PB-240 168

REMOTE SENSING OF POLLUTANTS

COMPUTERIZED REDUCTION OF LONG-PATH ABSORPTION DATA

National Oceanic and Atmospheric Administration

PREPARED FOR

National Environmental Research Center

July 1974

DISTRIBUTED BY:

NTIS
National Technical Information Service
U. S. DEPARTMENT OF COMMERCE

TECHNICAL REPORT DATA

(Please read Instructions on the reverse before completing)

1. REPORT NO.
EPA-650/2-74-113

2\.

PB 240 168

4. TITLE AND SUBTITLE
Remote Sensing of Pollutants: Computerized Reduction of Long-path Absorption Data

5. REPORT DATE
July 1974

6. PERFORMING ORGANIZATION CODE

7. AUTHOR(S)
V.E. Derr, M.H. Ackley, M.J. Post, and R.F. Calfee

8. PERFORMING ORGANIZATION REPORT NO.

9. PERFORMING ORGANIZATION NAME AND ADDRESS
Environmental Research Laboratories
National Oceanic and Atmospheric Administration
Boulder, Colorado 80302

10. PROGRAM ELEMENT NO.
1AA010; ROAP 26AAP

11. CONTRACT/GRANT NO.
EPA-IAG-077(D)

12. SPONSORING AGENCY NAME AND ADDRESS
National Environmental Research Center
U.S. Environmental Protection Agency
Research Triangle Park, North Carolina 27711

13. TYPE OF REPORT AND PERIOD COVERED
Final

14. SPONSORING AGENCY CODE

15. SUPPLEMENTARY NOTES

APTIC # 75063

16. ABSTRACT

Atmospheric gaseous pollutants are very numerous in industrial regions. It is estimated that 25 or more pollutant molecules may be found in the atmosphere in significant quantities. The measurement of the concentration of each gas from the complex spectrum obtained by a long-path infrared spectrophotometer requires the fitting of trial spectra composed from a library of spectra. The fitting procedure adjusts the concentrations of the trial spectra until a "best fit" in a least-squares sense is produced. This report is a description of the physical, mathematical, and calculational principles and procedures for the use of a digital computer program to determine concentrations of atmospheric gases in a path of a few kilometers. Detailed instructions for the computer program and a library of spectra are provided.

Reproduced by
NATIONAL TECHNICAL
INFORMATION SERVICE
US Department of Commerce
Springfield, VA. 22151

PRICES SUBJECT TO CHANGE

17. KEY WORDS AND DOCUMENT ANALYSIS

a. DESCRIPTORS	b. IDENTIFIERS/OPEN ENDED TERMS	c. COSATI Field/Group
Gaseous pollutants Infrared spectrophotometry Computer program		

18. DISTRIBUTION STATEMENT Unlimited	19. SECURITY CLASS *(This Report)* Unclassified	21. NO. OF PAGES 200
	20. SECURITY CLASS *(This page)* Unclassified	22. PRICE $7.25/2.25

EPA Form 2220-1 (9-73)

EPA-650/2-74-113

REMOTE SENSING OF POLLUTANTS

COMPUTERIZED REDUCTION OF LONG-PATH ABSORPTION DATA

by

V. E. Derr, M. H. Ackley,
M. J. Post, and R. F. Calfee

Wave Propagation Laboratory
Environmental Research Laboratories
NOAA
Boulder, CO 80302

IAG No. EPA-IAG-077(D)
ROAP No. 26AAP
Program Element No. 1AA010

EPA Project Officer: Dr. H. M. Barnes, Jr.

Chemistry and Physics Laboratory
National Environmental Research Center
Research Triangle Park, North Carolina 27711

Prepared for

OFFICE OF RESEARCH AND DEVELOPMENT
U.S. ENVIRONMENTAL PROTECTION AGENCY
WASHINGTON, D.C. 20460

July 1974

This report has been reviewed by the Office of Research and Development, Environmental Protection Agency, and approved for publication. Approval does not signify that the contents necessarily reflect the views and policies of the Environmental Protection Agency, nor does mention of trade names or commercial products constitute endorsement or recommendation for use.

Publication No. EPA-650/2-74-113

PREFACE

This report is arranged in a flexible digital format. The sections and subsections are designated by adding a decimal point and a number. For example, 5.1.2 is the second sub-subsection of the first subsection of section 5. References to sections are given without the use of the word section; for example, (7.3) refers the reader to the third subsection of section 7. Equations are invariably and uniquely referred to by a colon notation; that is, (3:7) refers the reader to equation seven of section 3. Figures and tables have a unique reference also; for example, F3.1 and T4.7 are respectively the first figure of section 3 and the seventh table of section 4. Appendixes are indicated by a capital A; that is, A2.1 is the first section of the second appendix. Figures, tables, and equations of appendixes are denoted respectively by FA3.3, TA2.4, and A3:2.

The International System of Units is used throughout except where the traditional unit still holds so strongly that confusion would result from absolute purity. Pollutants for example are often measured in parts per million. Conversion factors are given in A4.

This report is the product of a joint effort by EPA and NOAA. It conforms to EPA format rather than that of NOAA.

TABLE OF CONTENTS

Page

Abstract x

1. Introduction 1
  1.1 The Physical Problem 4
  1.2 The Mathematical Problem 9

2. Present Status of the Library of Absorption Spectra 19

3. Best-Fit Calculations of Concentrations From a Simulated Complex System 22
  3.1 Accuracy of Concentration Calculations 22
    3.1.1 Calculation of Concentrations for Spectra Without Noise 32
    3.1.2 Effect of Finite Signal-to-Noise Ratio (S/N) 33
    3.1.3 Effects of "Maverick" Gases in the Sample 34

4. Calculation of the Selected Gas Sample Concentrations From Spectrophotometer Data 39
  4.1 Effect of Mismatched Resolution of Experimental Data and Library Spectra 42
  4.2 Calculation of the Concentration of CO 50
  4.3 Calculation of the Concentration of SO_2 in the Presence of CH_4 56
  4.4 Calculation of the Concentration of CH_4 57

5. Conclusions and Recommendations 63

6. Acknowledgements 65

7. List of Symbols 66

8. References 67

A1. Sources and Contents of the Library of Spectra 71
  A1.1 Motivation for the Library 71
  A1.2 Sources and Content 71
  A1.3 Programs Used in Processing Spectra 91

Preceding page blank

Page

A2. Program EPAGAS Documentation and Analysis 111

A2.1 Program Summary 111

A2.1.1 List of Fortran Symbols and Descriptions 112

A2.1.2 List of Program and Subprogram Descriptions 121

A2.2 Simulation of Gaussian Noise 170

A2.3 Constituent and Transmission Error Formulation 171

A2.4 Interpolation Formulas Used in Transmission Library 173

A2.5 Program Execution Techniques 177

A2.5.1 Theoretical Gas Library Spectra 177

A2.5.2 Experimental Transmission Data 178

A2.5.3 Minimization of "Goodness of Fit" Controls 179

A2.5.4 Error Parameters 180

A2.5.5 Miscellaneous Execution Techniques 182

A3. Characteristics of the ROSE Long-Path Spectrophotometer System 185

A4. Convenient Conversion Factors 189

FIGURES

Page

F1.1(a-c) Infrared Spectra of Selected Atmospheric Constituents 11

F1.2 Infrared Spectrum of the Normal Atmosphere 14

F1.3 A Portion of the Transmission Spectra of CH_4, HNO_3, CH_4, and H_2O and Their Composite Spectra 15

F3.1 Principal Regions of Absorption of Selected Normal and Pollutant Atmospheric Constituents 30

F4.1 Spectrum of SO_2 and CH_4 Obtained by FTS (1300.8-1310.5 cm^{-1}) 40

F4.2 Spectrum of SO_2 and CH_4 Obtained by FTS (1310.6-1320.4 cm^{-1}) 41

F4.3 Spectrum of SO_2 and CH_4 Obtained by FTS (1320.5-1330.5 cm^{-1}) 41

F4.4 Comparison of a Theoretical Spectrum (High Resolution) of Water Vapor and the Spectrum Observed Through a Spectrometer With 1.0 cm^{-1} Slit Width 43

F4.5 Experimental Spectrum, CO, Nominal Resolution 2.5 cm^{-1} Obtained on ROSE System 44

F4.6 Calculated Spectrum, CO, Resolution 3.0 cm^{-1} 44

F4.7 Calculated Spectrum, CO, Resolution 1.0 cm^{-1} 45

F4.8 Calculated Spectrum, CO, Resolution 2.3 cm^{-1} 45

F4.9(a-c) Experimental Spectrum of Combination of SO_2 and CH_4, Nominal Slit Width 1.0 cm^{-1} (ROSE System) 47

F4.10(a-c) Calculated Spectrum, SO_2 and CH_4, Resolution 1.0 cm^{-1} 48

F4.11 Spectrum of CO Calculated by Program DEGRADE (Resolution 2.3 cm^{-1}) 52

F4.12 Spectrum of CO, Experimental Data From ROSE Spectrophotometer 52

F4.13 Computed Best-Fit Transmission of CO (EPAGAS Program), Resolution 3.0 cm^{-1} 54

F4.14 Computed Best-Fit Transmission of CO (EPAGAS Program), Resolution 2.3 cm^{-1} 55

F4.15 Change in Spectrum of CO for Small Change of Concentration (Calculated by Program DEGRADE) 56

F4.16 Computed Best-Fit Transmission of CH_4 Compared with the Given Experimental Transmission 60

FA1.1 Infrared Absorbing Regions of Atmospheric Gases 79

FA3.1 Resolution vs. Wavelength for Various Slit Width Settings for ROSE Spectrophotometer 187

TABLES

		Page
T1.1	Annual Average Concentrations of Important Pollutants in Cities (1968)	4
T2.1	Gases Available in the Spectral Library	21
T3.1	Calculation of Gas Concentrations From Simulated Composite Spectra (Set 1) (Infinite Signal-to-Noise Ratio)	23
T3.2	Calculation of Gas Concentrations From Simulated Composite Spectra (Set 2) (Infinite Signal-to-Noise Ratio)	24
T3.3(a-c)	Calculation of Gas Concentrations From Simulated Composite Spectra (Set 1) (S/N=10), (S/N=30), (S/N=100)	25
T3.4	Calculation of Gas Concentrations From Simulated Composite Spectra (Set 2) (S/N = 10, 30, 100)	28
T3.5(a-b)	Gas Concentrations From Simulated Composite Spectra (Set 1) (Maverick Gases)	36
T3.6	Gas Concentrations From Simulated Composite Spectra (Set 2) (Maverick Gases)	38
T4.1	Calculation of the Concentration of a Sample of CO by Application of EPAGAS to an Experimental Spectrum at Several Spectral Regions and Two Instrumental Resolutions	51
T4.2	Calculation of the Concentration of a Sample of SO_2 in the Presence of CH_4 by Application of EPAGAS to an Experimental Spectrum	58
T4.3	Calculation of the Concentration of a Sample of CH_4 and H_2O by Application of EPAGAS to an Experimental Spectrum	62
TA1.1	Spectra From the WPL Data Set	72
TA1.2	Spectra Processed From the Scientific Literature	75

ABSTRACT

Atmospheric gaseous pollutants are very numerous in industrial regions. It is estimated that 25 or more pollutant molecules may be found in the atmosphere in significant quantities. The measurement of the concentration of each gas from the complex spectrum obtained by a long-path infrared spectrophotometer requires the fitting of trial spectra composed from a library of spectra. The fitting procedure adjusts the concentrations of the trial spectra until a "best fit" in a least-squares sense is produced. This report is a description of the physical, mathematical, and calculational principles and procedures for the use of a digital computer program to determine concentrations of atmospheric gases in a path of a few kilometers. Detailed instructions for the computer program and a library of spectra are provided.

1. INTRODUCTION

The Chemistry and Physics Laboratory of the Environmental Protection Agency (EPA) in Raleigh, NC, has been studying the feasibility of various methods of measuring the gaseous pollutants of the earth's atmosphere. One of the most important and promising of these methods is the use of long-path infrared absorption by gases of the atmosphere. This method, useful only for gas concentration measurement and not for measuring the concentration of aerosols, employs a beam of radiation of wide bandwidth from a stable source strong in infrared radiation. The collimated beam passes through a part of the atmosphere to be studied, typically about 1 to 4 km. It may be received at the end of that path or returned by reflectors to a receiver near the transmitter. The receiver contains an infrared spectrometer.

Many methods other than long-path infrared absorption have been considered for the analysis of atmospheric gases. Much effort has been expended to measure accurately the concentrations of pollutants near cities and in the ambient atmosphere far from cities to obtain background concentrations. Other methods include Raman spectroscopy, the use of tunable laser sources, and observation of the ultraviolet electronic spectra of molecules. These methods may either provide the entire analysis or contribute partially. However, in this report, we confine our attention to the use of long-path infrared absorption spectra and attempt to exploit this method as fully as time and resources allow. It is the authors' belief that other methods will prove important in the future, but the long-established science of infrared absorption spectroscopy allows important progress to be made very quickly so that at least preliminary data may be obtained on the pollutants in the atmospheres of cities. It is probable that combinations of several methods will be utilized.

The analysis of the composite infrared spectra prduced by many atmospheric constituents is laborious and difficult. It requires a large computer and a library of spectra. The Wave Propagation Laboratory of the National Oceanic and Atmospheric Administration (NOAA) in Boulder, CO, has worked for many years in a joint program with EPA to aid in the analysis of such composite spectra. The EPA financed portion of the work has been done under contract number EPA-IAG-077(D) entitled, "Remote Sensing of Pollutants."

This report is specialized in several ways. Firstly, it makes the assumption that the instrument used in obtaining the data is that described in 1.1. This assumption is necessary in order to be specific concerning the parameters used in our computational method. Without such an assumption, it would be necessary to consider wide ranges of parameters of a long-path spectrophotometer. The complexity caused by this would not clarify the method and would complicate the analysis. It is anticipated that this specialization will not hamper the extension of the method to other specific devices. This has been accomplished by letting the input parameters to the computer program which analyzes the complex composite spectra to be at the discretion of the user. Thus, the choice of a different instrument would merely be reflected in different input parameters to the calculation. A second way in which the work has been limited is by the choice of spectra placed in the library. This choice has been dictated by the impossibility of obtaining a complete library of spectra with the resources available.

Fortunately, for the immediate application of the method, a relatively small library of spectra is probably quite adequate. The spectra of the gases of importance in the Los Angeles, CA, atmosphere, the Raleigh, NC, atmosphere, and the atmosphere near refineries have been given preference whenever possible. However, the library of spectra has been limited by the unavailability of suitable spectra of some gases. Spectra were obtained from the scientific literature whenever possible. In certain cases, spectra were determined by laboratory experiment when they were not available in the literature. The entries and sources of the spectral library are given in A1.

The major normal atmospheric gases, H_2O, CO_2, N_2O, and CH_4, are part of the population. (Oxygen and nitrogen have no significant spectra in the frequency range of interest.) Normal constituents and pollutants are discussed in 2. The listed gases were chosen because of their presence in the atmosphere and because detailed spectroscopic parameters are available for them, including either line positions, strengths, and line-broadening constants, or high-resolution experimental spectra.

The spectra of molecules have long been used as a definitive identifier, and when the spectrophotometer is properly calibrated, molecular concentrations can also be measured. In many cases, however, the method has been applied to relatively large concentrations in the laboratory, to pure gases, or at most to a mixture of a few constituents. It is true that laboratory spectrometers have been used to obtain low-level concentrations of a few species in the midst of large concentrations of other gases, but this method can only be successfully performed by simple observation when the regions of significant spectral absorption are well separated. Because most organic pollutant molecules have rich infrared spectra, this spectral region may be used for detection and measurement, but the potential for interference between overlapping spectral lines is high.

Thus, the measurement of atmospheric pollutants is complicated by low concentrations and by the simultaneous presence of many gases in temporally and geographically varying amounts. Because low concentrations may be significant for biological well-being, the method of analysis must be sensitive. Table 1.1 (Burriss et al., 1972)* shows typical concentrations of some pollutants in several typical cities. The sensitivity must be sufficient to determine such concentrations and, perhaps, a factor of 10 less. The interference due to aerosols is not considered in this report.

Because hydrocarbons (HC) are a large percentage of the total atmospheric pollutant content of many cities, they need special attention. In Chicago, IL, in 1968, the average HC concentration is 5380 $\mu g/m^3$, compared with CO at 7750 $\mu g/m^3$. The molecules making up the total HC content of the atmosphere for any location are not clearly known because many current analyzes separate only methane and non-methane HC. However, if there is no outstanding single HC and the concentration of HC is distributed among many different molecules, the weak absorption of each will provide

---

* References are to authors and year of publication.

Table 1.1

ANNUAL AVERAGE CONCENTRATIONS
OF IMPORTANT POLLUTANTS IN CITIES (1968)

City	CO	HC*	NO_2	NO	SO_2
Chicago, IL	7750	5380	98	94	340
Cincinnatti, OH	7000	4220	59	-	57
Philadelphia, PA	-	2690	78	67	230
Denver, CO	6750	5380	78	54	29
St. Louis, MO	5750	7300	39	40	86
Washington, DC	4250	2690	98	54	110

*HC = hydrocarbons

Units: $\mu g/m^3$

(From Burriss et al., 1972)

effectively a small absorption background which can be ignored. This problem can be solved definitively only after a detailed analysis of constituents is obtained for a particular region, and the spectra of these constituents are known.

Thus, the considerations discussed above require the analysis of a complex spectrum of many gases of relatively low amounts. This report is a description of the physical, mathematical, and computational procedures necessary to obtain the concentrations of many gases contained in an atmospheric sample by long-path infrared spectroscopy. The types of gases present in the sample are assumed known; the concentrations are unknown.

1.1 THE PHYSICAL PROBLEM

We consider an atmosphere composed, perhaps inhomogeneously, of normal and pollutant constituents. The pressure is assumed to be one atmosphere (1.02×10^6 dynes/cm^2). We assume that the constituents do not significantly change during the course of an observation. The observation is through a portion of the atmosphere approximately 61 cm in diameter and 1 to 4 km long. (These dimensions arise from the characteristics of a long-path spectrophotometer (see A3).) To focus on a specific problem, we choose the Research Triangle Area of North Carolina and a representative concentration of pollutants for that area will be used when available. For those pollutants whose average concentrations in Raleigh are not known, the known values in other comparable areas will be used. If no reliable values of trace materials can be obtained, they will be assumed to lie in the few parts per billion range. Further information may be found in A1, where TA1.1 and TA1.2 show the lowest concentrations assumed.

Although the method to be described is quite general, for clarity we will use the parameters of a specific system for remote sensing of pollutants (Streiff and Claysmith, 1972). This system, when used in an absorption mode, is a field-type infrared (IR) scanning spectrophotometer with a blackbody light source collimated by a telescope for long-path operation. Some system parameters of interest are given in A3.

The equipment can be used as a receiving spectrophotometer to study emissions from any sufficiently hot source, since it was designed to measure the emission spectra of gases in smokestacks. The methods developed herein are applicable to the study of emissions by substituting a library of emission spectra in place of the library of absorption spectra.

Thus, by means of the long-path spectrophotometer, we obtain the superimposed spectra of the atmospheric constituents in the spectral ranges 3 to 5 μm (3330-1820 cm^{-1}) and 7 to 13.5 μm (1430-740 cm^{-1}). The 5 to 7 μm spectral range is not used because of the very strong absorption of water vapor. It is the purpose of the method described in this report to obtain from the superimposed spectra the concentrations of the molecular species present in the path.

Each of the gases present in the path acts independently of the others according to the Bouguer-Beers law. That is, a small thickness of gas absorbs a fractional amount, $\frac{dI}{I} = -K'dx$, which implies

$$I = I_o e^{-K'x} \quad , \qquad (1:1)$$

where I_o is the initial intensity, K' is the absorption coefficient, and x is the absorber thickness. (Physically, K' is the reciprocal of the distance at which the intensity is 1/e of the initial intensity.) If the first absorber is followed by a second, equal to the first, it absorbs an equal fraction of the radiation incident upon it. If we consider a plane wave impinging on "slabs" of absorbing gases, the effect is the same whether the slabs follow one another or the gases are all placed in one slab. If the gases have absorption coefficients K_1', K_2', etc, then the intensity after passage through the absorbers is $I = I_o e^{-(K_1'+K_2'+\ldots)x}$. The quantities K_1', K_2', ... are complicated functions of the frequency (herein designated as ν and measured in cm^{-1}) and of the temperature, pressure, and concentrations of the gases. Hence I is also a complicated function of the same variables. The long-path spectrophotometer measures I/I_o as a function of frequency and, from that data, it is necessary to determine the concentration of the gases.

To see how the concentration enters the problem, we write a K'x as $K\rho x = KW = \sigma N_o x$, where N_o is the number of absorbing molecules per unit volume, ρ is the density of absorbing gas per unit area of the radiation path, σ is the molecular absorption coefficient, W is the area density of gas, and K is the absorption coefficient. A very wide range of units has been used for these quantities in the literature. The reader should refer to Deutschman and Calfee (1967) and Calfee (1971) for transformations between commonly used sets of units. Table A1.4 lists a summary of these transformations. In this report, W is in units of molecules per square centimeter and I is in units of centimeters squared per molecule.

If the spectrophotometer measures I/I_o as a function of wavelength after suitable adjustment and calibration, the data available (by taking the natural logarithm of each side of 1:1) is the product $KW = \ell n\frac{I}{I_o}$, where KW is a function of frequency. But $KW = \sum_{i=1}^{N} K_i W_i$, where N is the total number of absorbing gases, K_i is the absorption coefficient, and W_i is the (area) density of the i^{th} gas. If we have the necessary spectral information about the gases, the K_i is known, and we need to determine the concentrations (determinable from W_i) of the gases present. The ROSE* system allows for an I/I_o output by having an internal blackbody that can be set at the same temperature as the remote blackbody. Their signals fall on the same detector with different chopping frequencies. They are separated and amplified by two phase sensitive locking amplifiers, and then ratioed. The problem that occurs (spurious variation of I/I_o) is presumably due to the differences in the optical paths of the two beams. Thus a true I/I_o is not provided.

If the instrument only provides relative transmission, I, rather than I/I_o, it is necessary to scale the data before the calculation of concentrations. The scaling must be done so that the input data to the computer program lies between zero and one, just as does I/I_o. This can be accomplished by simple scaling only if the response of the spectrophotometer is the same over all frequency ranges. If not, the data must be scaled in sections. In either case, only relative concentrations can be obtained. If absolute concentrations are required, and I/I_o is not provided by the spectrophotometer, they may be obtained by adding the spectrum of a known amount of calibrating gas. There are several ways of doing this in practice. If, for example, the humidity is known along the path, this could provide an absolute scale. The accuracy would suffer, however, for obtaining low-level concentrations in comparison with the relatively larger concentration of water vapor, and it would be more accurate to choose an independent measurement of a gas of lower concentration. Another method would be to insert a test cell containing a gas of known concentration, adjusted so that its absorption falls in a spectral range unlikely to interfere with the unknown spectra. The concentration should be adjusted so as to fall within the range of most of the unknown gases, by trial and error, if necessary. The material of the cell must be nonabsorbent or accounted for by observing the effect of an empty cell. I_o can also be obtained by recording a spectrum upwind of an extended source or in a clean environment. The quantity can then be obtained downwind at the same path length and reference blackbody temperature. The resulting I/I_o will then be an absolute spectrum.

---

* Remote Optical Sensing of Emissions

The problem presented by the infrared continuum can under some circumstances be obviated by this method. The infrared continuum, a general increase of all absorption curves over a large part of the spectral range being considered here, occurs when long paths though the atmosphere are used. At present, only empirical methods of dealing with it are known. Over the relatively short path (2-4 km) planned for the spectrophotometer described in this report, the problem will often arise. (See Shapiro and Gush (1966), Burch (1970), and McClatchey et al., (1971).) However, in contamination such as occurs in Los Angeles, the continuum is often a significant contribution to the spectrum.

Determining concentrations is complicated by the inescapably finite signal-to-noise ratio (S/N) of the spectrophotometer and the noise due to atmospheric scintillation. Thus, the signal from the spectrophotometer has a noise component. It is probably not additive, but has been nonlinearly mixed by the system detector. However, a complete noise analysis of the detector is not available, and the character of the noise is dependent on the particular system. Therefore, we will make an assumption, for convenience, that the noise on the signal is additive Gaussian noise, that is, the output of the system is composed of a signal and noise, (I/I_o+N), where the mean value of N is zero. The bandwidth of the Gaussian noise is determined by the time constant of the circuitry; it has essentially the same spectral bounds as the signal I/I_o. The precision of the measurement is dependent on the S/N ratio and it, together with the resolution of the ratio digital voltmeter, determine the error. The ratio digital voltmeter of our system has a resolution of 1 part in 10^4. However, the S/N ratio depends on atmospheric turbulence and is probably never as large as 10^4 (in the absorption mode). Hence the signal-to-noise ratio is usually the limiting factor in the precision of concentration measurement.

Streiff and Claysmith (1972) discuss the signal and noise characteristics of the spectrophotometer system. Their data and a private communication from M. Streiff indicate that practical working ranges of S/N run from 10 to 100. Without further tests of the system under working conditions, a complete noise analysis cannot be performed. To offer guidance on the degradation suffered in precision under finite signal-to-noise ratios, the computer analysis has been performed with an added simulated noise (see section 3.1.2).

In a different sense than the noise of atmospheric fluctuations and the spectrophotometer system discussed above, the presence of water vapor and to some extent that of natural carbon dioxide in the atmosphere constitute important sources of noise or interference. Water vapor concentration varies considerably in time and space, and its very strong absorption spectrum is found through most of the frequency region of interest. It interferes least with other spectra in the region from 2400 to 2600 cm^{-1}. Unfortunately, only a few gases of interest have observable absorptions there, and thus water vapor must always be one of the unknown gases whose concentration is determined by the procedure described herein. If it is measurable otherwise, it may be desirable to consider removing its effect from the data by preprocessing. When the absolute humidity and temperature of the sample are known, the water vapor spectrum may be computed (1.2) and divided from the absorption spectrum to eliminate its effect. Carbon dioxide is also a naturally variable constituent, but its variation and absorption are less than those of water; and it usually is not an important interferant.

1.2 THE MATHEMATICAL PROBLEM

The transmission spectrum of an atmospheric constituent I/I_o is a unique function of frequency ν, temperature T, pressure p, and concentration W, that is, $(I/I_o)_i = f_i(\nu, W_i, P, T)$, where the subscript indexes the specific gas. Figure 1.1 (a-c) shows typical spectra of normal constituents and pollutants, F1.2 shows the normal atmosphere. If the transmission path contains two different gases, the resulting transmission spectrum is obtained by application of the Bouguer-Beers law (1:1). According to the discussion of 1.1, the spectrum due to two gases is obtained by multiplying the relative transmissions, that is,

$$(I/I_o) = (I/I_o)_1 \; (I/I_o)_2 = e^{-(K_1W_1+K_2W_2)} \; , \qquad (1:2)$$

and in general, for N gases,

$$(I/I_o) = \prod_{i=1}^{N} (I/I_o)_i = e^{-\sum_{i=1}^{N}(K_iW_i)} , \qquad (1:3)$$

where (I/I_o) is the transmission of the composite atmosphere.

It should be noted that in whatever part of the frequency range one of the gases has zero (or small) transmission, the composite spectrum (I/I_o) is also zero. In any part of the frequency range where the transmission is one (no absorption) for all gases but one, the spectrum will be identical to that one. Otherwise, the spectrum will be the product of the constituent spectra and will generally present a complex problem to the experimenter attempting to identify the constituents and measure their concentrations. Figure 1.3 shows spectra of two gases, CH_4 (methane) and HNO_3(nitric acid), and their combined spectrum.

The spectra of the gases $(I/I_o)_i$ are functions of ν, T, P, and W. We assume that only the concentrations W_i are unknown, and the compound spectrum recorded as a function of frequency ν is the output of the spectrophotometer system.

In practice, $(I/I_o)_i$ as found in the library is accurate for only those values of P and T for which the spectrum was determined. Departures of $(I/I_o)_i$ from true (I/I_o), however, are negligible for the range of temperatures and pressures found on the earth's surface.

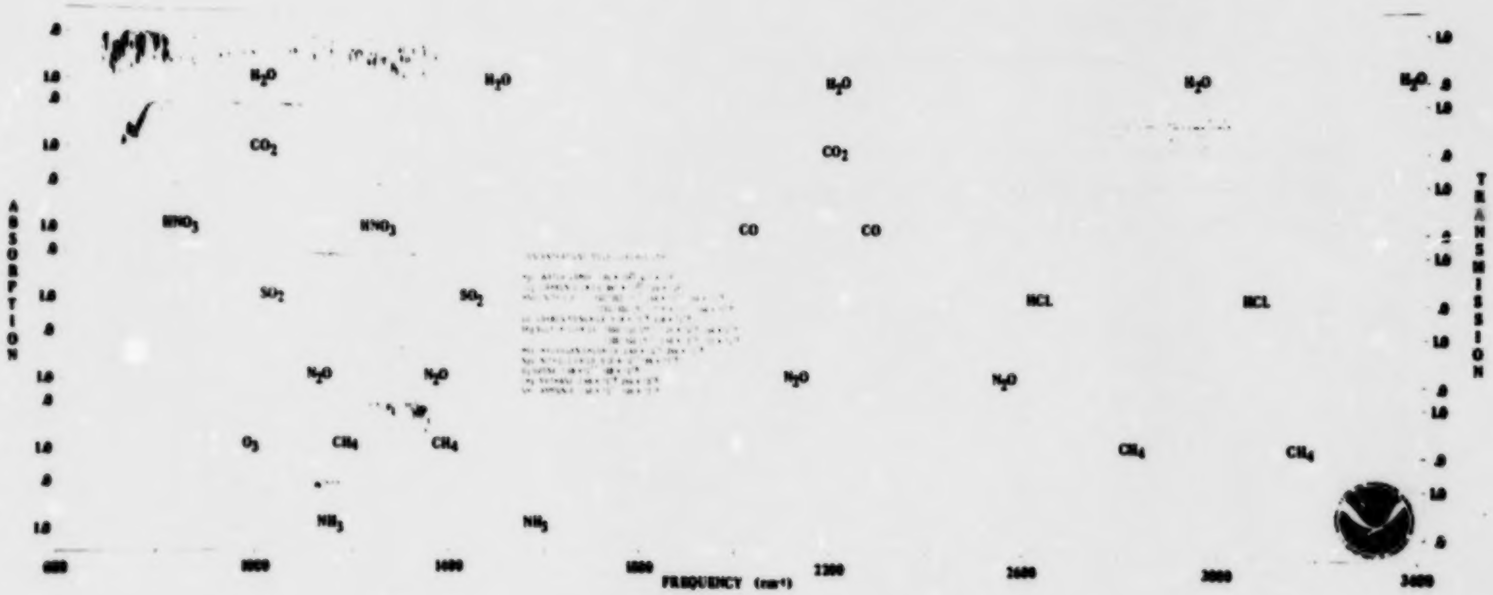


Figure 1.1a

INFRARED SPECTRA OF SELECTED ATMOSPHERIC CONSTITUENTS
(Bands are identified on the overlay)

Figure 1.1b
INFRARED SPECTRA OF SELECTED ATMOSPHERIC CONSTITUENTS

Reproduced from best available copy.

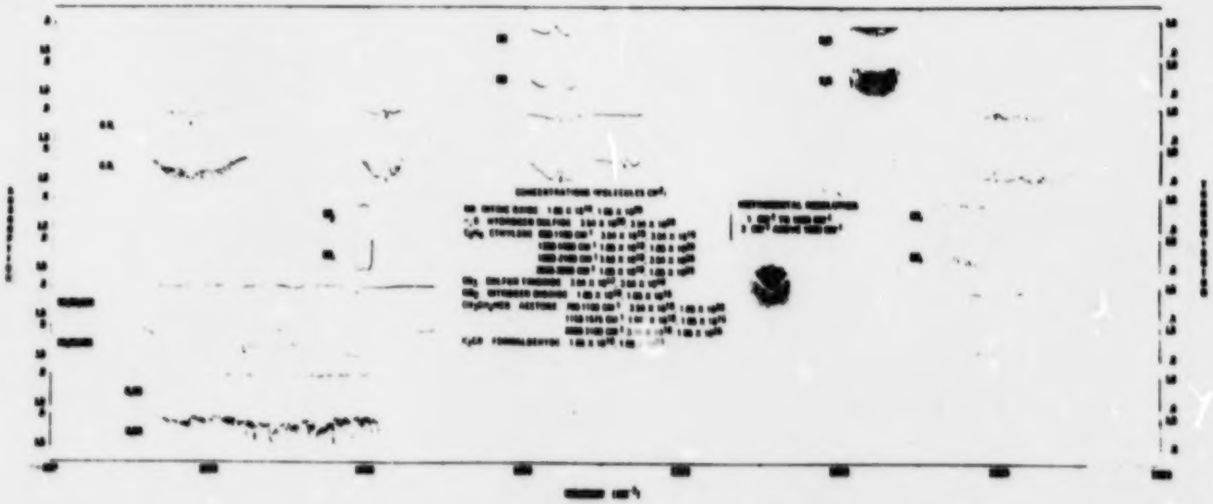


Figure 1.1c
INFRARED SPECTRA OF SELECTED ATMOSPHERIC CONSTITUENTS

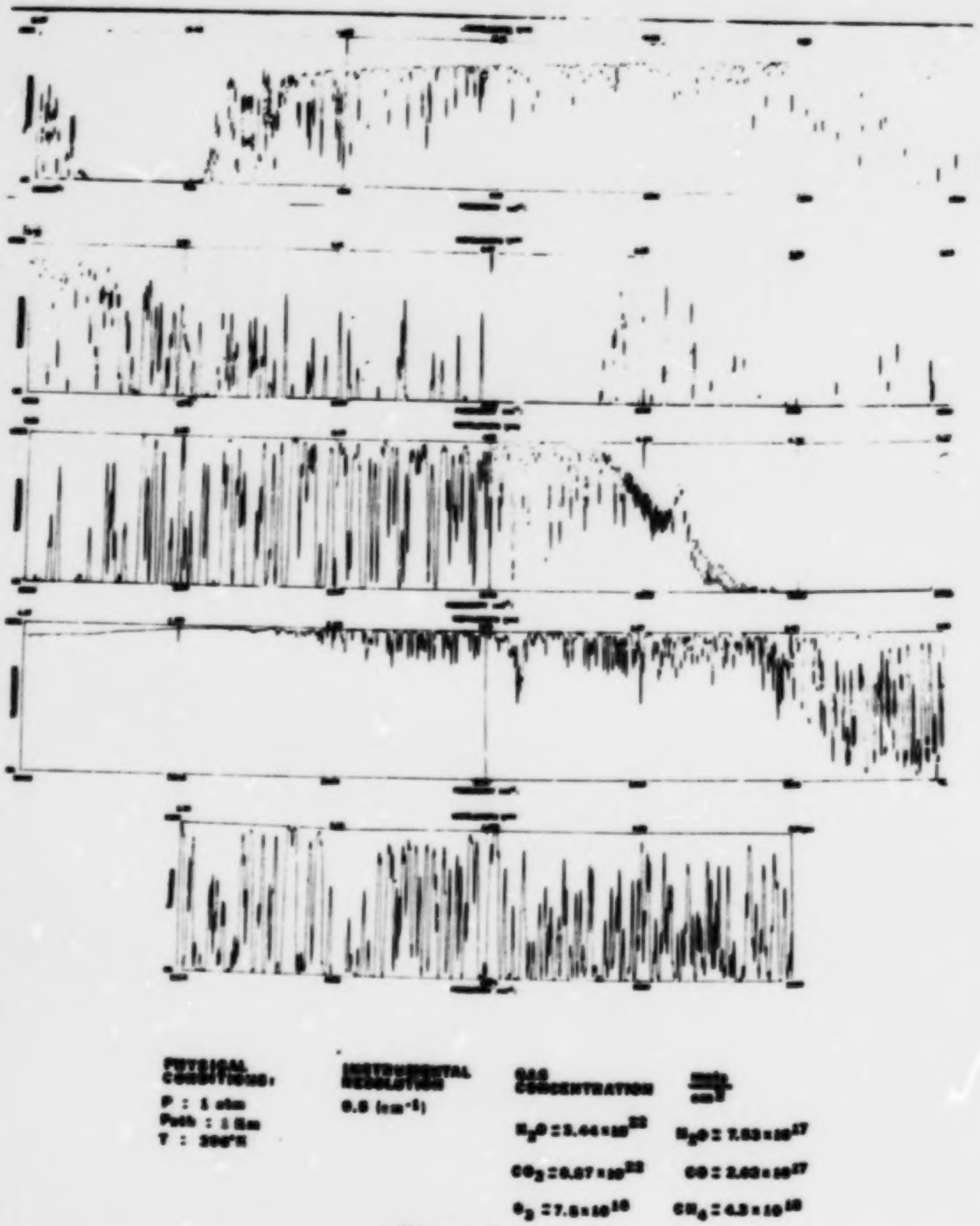


Figure 1.2

INFRARED SPECTRUM OF THE NORMAL ATMOSPHERE

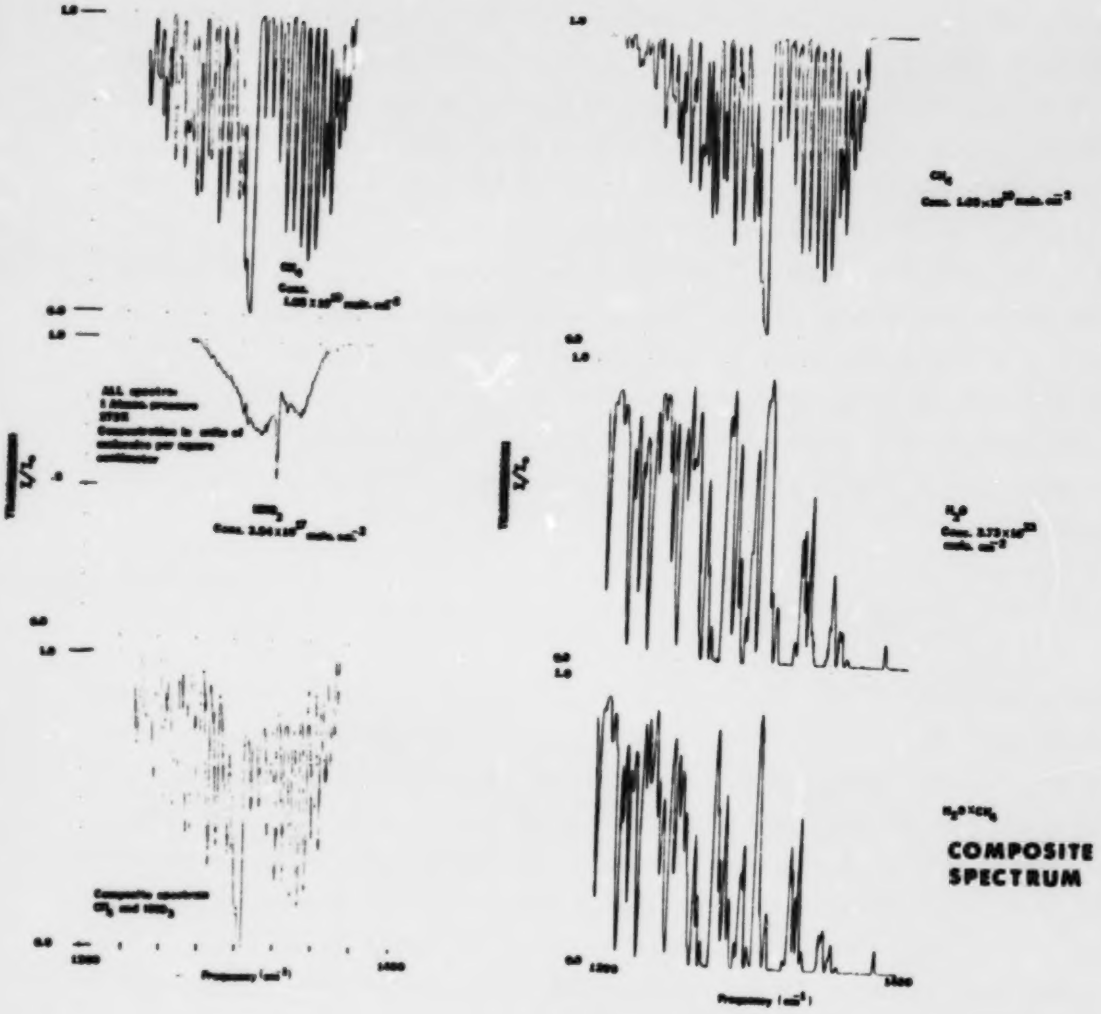


Figure 1.3
A PORTION OF THE TRANSMISSION SPECTRA OF
CH_4, HNO_3, CH_4, AND H_2O AND THEIR COMPOSITE SPECTRA

Therefore, in this application p and T are assumed constants, with $p = 1$ atm(1.02×10^6 dynes/cm^2) and $T = 296^{\circ}$K. For special circumstances, accurate $(I/I_o)_i$ may be calculated for any temperature and pressure only for molecules whose line parameters are fully known. This is accomplished using program DEGRADE (Deutschman and Calfee, 1967). For spectra taken experimentally, there is no known transformation to a new set of P and T without a detailed study of the molecule.

We call the output of the spectrophotometer $S(\nu)$. We must choose the concentrations W_i for all the gases which are in the sample path in such a way that the spectrum (I/I_o) of (1:3) is the same as, or as close as possible to, the experimental data $S(\nu)$. Simply put, this is performed by choosing a trial set of W_i, calculating (I/I_o), and then comparing (I/I_o) with $S(\nu)$. The comparison is made by calculating an error function:

$$E_j = \left\{ \frac{1}{n} \sum_{\nu}^{n}{}, \left[(I/I_o;\ W_i^j) - S(\nu) \right]^2 \right\}^{\frac{1}{2}} . \qquad 1{:}4$$

Here the symbol $(I/I_o;\ W_i^j)$ signifies the composite transmission spectrum, calculated from the spectral library under the assumption that the ith gaseous constituent has the trial concentration W_i^j . Of course $(I/I_o;W_i^j)$ is a function of frequency. The number of terms in the sum is n. The summation over ν is somewhat arbitrary; maximum accuracy will be achieved by using all statistically independent measurements. A good compromise for the number of points was found by trial and error to be 5 per resolution element. (In the calculation described in this report, the data $S(\nu)$ and the spectra of the individual gases are stored at increments of $\Delta\nu = 0.2$ or 0.6 cm^{-1}, depending on the frequency range (see section 2).)

The trial spectra of individual gases are obtained by interpolation and extrapolation from a set of spectra given in the library for specific concentrations which span the range of concentrations expected in the atmospheric sample.

The first estimate of the concentrations, W_i^j, is generally not correct, of course. After the calculation of the first error, E_1, the concentrations are varied in an orderly stepwise manner, calculating successive corresponding errors E_j, until the set of concentrations is obtained that minimize E_j. These concentrations are then the "best" estimate of the concentrations in the atmospheric path in which the sample data $S(\nu)$ were obtained.

The calculation is performed by the computer subroutine MINMYZD, developed by Slutz and Winkelman (1964) and applied by Lawrence and Hallenbeck (1965). MINMYZD is described in detail in A2.

This computer routine, although in principle as simple as described above, requires careful design to prevent costly, long computer runs and inaccuracies. The problem is to find the minimum value of a multidimensional function, E, whose variables are the W_i. Care must be taken to avoid local minima in this function and to find the absolute minimum. The subroutine MINMYZD has features which practically assure that result in applications such as herein discussed.

It is not possible and not desirable to discuss here the problems of uniqueness involved in minimization problems such as this one. Instead, the long experience with this program and its successful application to many physical problems gives assurance of its usefulness and accuracy.

Arbitrarily chosen initial concentrations prolong the computer running time and may cause inaccuracy in the final results. Hence initial concentrations are selected for each constituent by computing a "best fit" in a very limited spectral range where the constituent has a prominent spectrum and the interference from other gases is small. These selected concentrations are then employed as initial, approximate choices in the computation over the whole spectral range. When the number of gases is small, these initial values may be sufficient for the experimenter's needs.

It is true that the optimization or best-fit computation here employed is not the only way the concentrations could be obtained from the experimental data. Many other methods exist, under the names of best-fit, minimization, or inversion methods. No attempt was made to evaluate which of these would be the best method. The MINIMYZD routine was selected because of its long history of successful efficient operation. Methods such as that of Backus and Gilbert (Westwater and Cohen, 1973) offer some advantages of error estimation. Matthes (1973) has devised a method to circumvent problems arising in least-squares fitting which results in negative concentration estimates. This problem may be prevented in the present program by placing a lower bound of zero on any concentration estimate, but it is probably better practice to eliminate such absolute constraints on the MINMYZD subroutine to prevent nonlinear effects on the search process.

2. PRESENT STATUS OF THE LIBRARY OF ABSORPTION SPECTRA

It is essential in the identification of pollutant gases by long-path absorption spectroscopy to possess accurate spectra of the gases expected to occur in the atmosphere under examination. These spectra must be complete in the spectral range covered by the long-path spectrophotometer because "maverick" gases (i.e., gases present in the sample but not in the spectral library) cause inaccuracy in the calculation of the concentrations. Some examples of such inaccuracies are presented in 3.

The spectra must be known in such a way that they may be determined as a function of temperature, pressure, and concentrations to suit the conditions in the test atmosphere. If, for example, the line position, the line strength, and the broadening parameters of a gas as functions of temperature and pressure are known, the spectra may be computed for any concentration by computer routines (Deutschman and Calfee, 1967, and McClatchey et al., 1973). If these parameters are unknown, but accurately measured experimental spectra are available for concentrations near those of interest, we may extrapolate from them to other concentrations.

To make use of such experimental data, it is necessary that the spectra be obtained at a temperature and pressure near those normally encountered over polluted cities, that resolution is comparable to or better than that of the spectrophotometer used, and that the ordinate of the spectra is calibrated as an absolute transmittance (or, that the line strengths in some way are known).

Whether such spectra are available depends on whether researchers have been interested in these specific molecules. No concerted effort has been made to the authors' knowledge to obtain detailed spectra of all atmospheric constituents including pollutants. However, an effort has been made by a loosely knit group of researchers* to obtain detailed spectra on the normal atmospheric constituents. These data are summarized by McClatchey et al., (1973).

---

* The Group on Atmospheric Transmission Studies (GOATS).

In general, however, the search of literature on the spectra of normal and pollutant atmospheric constituents reveals a wasteland punctuated by a few oases of excellent work. Spectroscopists, academic and industrial, have not been supported in obtaining the high-resolution (1 cm^{-1}) infrared spectra with a well calibrated intensity scale at atmospheric pressure needed for concentration measurements in an atmosphere containing many interfering gases.

The chief limitation of the application of this (or any) spectroscopic technique to the detection of atmospheric constituents lies in the ignorance of the spectra of many atmospheric constituents, particularly pollutants. Just as the criminal's fingerprints cannot be used to identify him unless they are in police files, so observation of a gas spectrum only confuses as if it is not in our library. Notably, spectra are unavailable for many hydrocarbons. Also spectra of excited (high temperature) molecules such as found in smoke stacks are not in the literature.

The sources of the spectral library are summarized in A1 and detailed descriptions are given there of the parameters of the spectra. Also described are manipulations performed in converting spectra found in the literature into computer compatible form. We summarize in T2.1 the gases available in the spectra library. Spectra for selected concentrations for all these gases are displayed in F1.1(a-c).

Table 2.1
GASES AVAILABLE IN THE SPECTRAL LIBRARY

Gas	Formula	Accuracy of Spectrum *
Water	H_2O	1
Carbon dioxide	CO_2	1
Ozone	O_3	1
Nitrous oxide	N_2O	1
Methane	CH_4	1
Carbon monoxide	CO	1
Hydrochloric acid	HCL	1
Sulfur dioxide	SO_2	2
Nitric acid	HNO_3	2
Ammonia	NH_3	2
Nitric oxide	NO	1
Hydrogen sulfide	H_2S	2
Ethylene	C_2H_4	2
Nitrogen dioxide	NO_2	2
Acetone	CH_3CH_2HCO	3
Formaldehyde	H_2CO	2
Ethane	C_2H_6	2
Normal-butane	$n\text{-}C_4H_{10}$	3
Iso-butane	$i\text{-}C_4H_{10}$	3
Propane	C_3H_8	3
Pentane	C_5H_{12}	3

*Accuracy of spectrum:
(1) Line information highly trustworthy;
(2) Good experimental spectrum available, trustworthy; and
(3) Low resolution only, some uncertainty.

3. BEST-FIT CALCULATIONS OF CONCENTRATIONS FROM A SIMULATED COMPLEX SYSTEM

The application of the MINMYZD fitting routine to specific problems requires a controlling program to direct progress from the input of data through the minimization process to the final gas constituent concentrations and error estimates. This program, called EPAGAS, is described in detail in A2.

Using this program, the entire computation was tested by preparing trial spectra, simulating the data obtained from a long atmospheric path spectrophotometer experiment, and then obtaining the concentrations. The tests were made "blind", that is, the programmer was presented with the simulated sample data without knowledge of the concentrations until the computer results were compared with the known composition. In this section, a number of such computations are reviewed. The objectives were to "debug" the program, to test the success of determining the concentrations of gases from their composite spectrum, to determine the accuracy achievable, and to estimate the degradation of accuracy due to finite signal-to-noise ratio in the data. Further, the effects of a lack of a one-to-one correspondence of the library of spectra with the atmospheric constituents are examined.

In these preliminary tests, library spectra of H_2O, N_2O, CH_4, SO_2, HNO_3, CO, CO_2, and O_3 were employed in various concentrations of combinations to produce a complex spectrum.

3.1 ACCURACY OF CONCENTRATION CALCULATIONS

Tables 3.1 through 3.4 list the trial runs made on the combinations of gases indicated at the top of the table. Each fitting was made over a narrow frequency range to determine the suitability of that range for detection of the gas under investigation and to observe inaccuracies arising when the gas has no absorption in the region.

It should be noted here, and it is emphasized and clarified in A2, that the analysis of complex spectra by the method which is the subject of this report should not be done blindly or automatically.

Table 3.1

CALCULATION OF GAS CONCENTRATIONS FROM SIMULATED COMPOSITE SPECTRA (SET 1) (INFINITE SIGNAL-TO-NOISE RATIO)

Code	Wavenumber Range (cm^{-1})	H_2O (1A1) Wg = 3.73×10^{22}		N_2O (4A1) Wg = 1.08×10^{18}		CH_4 (5A1) Wg = 1.08×10^{19}		SO_2 (20.2) Wg = 3.54×10^{17}		HNO_3 (30.2) Wg = 3.54×10^{17}		Final Transmission Error	
		Wc	ΔW	Wc	ΔW	Wc	ΔW	Wc	ΔW	Wc	ΔW	RMS Wtd.	Unwtd.
8868 a	1230.0-1249.8	3.73×10^{22}	.00	1.09×10^{18}	-.01	1.08×10^{19}	.00	17.88×10^{17}	-14.34	0.07×10^{17}	3.47	.03	6.60
b	1250.0-1269.8	3.73×10^{22}	.00	1.09×10^{18}	-.01	1.08×10^{19}	.00	17.88×10^{17}	-14.34	0.10×10^{17}	3.41	.06	6.60
c	1270.0-1289.8	3.73×10^{22}	.00	1.08×10^{18}	.00	1.08×10^{19}	.00	17.88×10^{17}	-14.34	3.55×10^{17}	- 0.01	.03	6.41
d	1290.0-1309.8	3.73×10^{22}	.00	1.10×10^{18}	-.02	1.08×10^{19}	.00	0.75×10^{17}	2.79	3.45×10^{17}	0.09	.05	1.25
e	1310.0-1329.8	3.73×10^{22}	.00	1.23×10^{18}	-.15	1.06×10^{19}	.02	4.29×10^{17}	- 0.75	3.42×10^{17}	0.12	.03	0.35
f	1330.0-1349.8	3.73×10^{22}	.00	1.21×10^{18}	-.13	1.08×10^{19}	.00	3.51×10^{17}	0.03	3.56×10^{17}	- 0.02	.02	0.06
g	1350.0-1369.8	3.73×10^{22}	.00	1.21×10^{18}	-.13	1.06×10^{19}	.02	3.32×10^{17}	0.22	4.32×10^{17}	- 0.78	.03	0.37
h	1370.0-1389.8	3.68×10^{22}	.05	1.21×10^{18}	-.13	1.48×10^{19}	-.40	4.88×10^{17}	- 1.34	18.38×10^{17}	-14.84	.03	6.68
i	1380.2-1400.0	3.67×10^{22}	.06	1.21×10^{18}	-.13	1.48×10^{19}	-.40	5.10×10^{17}	- 1.56	18.23×10^{17}	-14.69	.03	6.61

Wg = given concentration in simulated sample, molecules cm^{-2}

Wc = calculated concentration, molecules cm^{-2}

ΔW = Wg-Wc (the factor of a power of ten is omitted), molecules cm^{-2}

See 3.1 for definition of the weighted and unweighted final transmission errors (last two columns). Numbers following gas formulae are computer codes.

A composite sample spectrum was obtained by adding the absorption coefficients as a function of frequency and calculating the transmission (See 1.2). The sample spectrum was then analyzed by MINMYZD to obtain the calculated concentrations. The error is the difference between the calculated and given transmission. (See 3.1).

Table 3.2

CALCULATION OF GAS CONCENTRATIONS
FROM SIMULATED COMPOSITE SPECTRA (SET 2)
(INFINITE SIGNAL-TO-NOISE RATIO)

Code	Wavenumber Range (cm^{-1})	H_2O (1B1) Wg = 3.73 x 10^{21}		CO_2 (2B1) Wg = 1.08 x 10^{21}		O_3 (3B1) Wg = 2.69 x 10^{17}		N_2O (4B1) Wg = 1.08 x 10^{18}		CO (6B1) Wg = 1.88 x 10^{19}		Final Transmission Error	
		Wc	ΔW	Wc	ΔW	Wc	ΔW	Wc	ΔW	Wc	ΔW	RMS Wtd.	Unwtd.
9999 a	2100.0-2159.4	$3.74x10^{21}$	-.01	$1.08x10^{21}$	.00	$2.84x10^{17}$	-0.15	$1.22x10^{18}$	-.14	$1.88x10^{19}$	0.00	.03	0.09
b	2160.0-2219.4	$3.72x10^{21}$	.01	$1.05x10^{21}$	.03	$5.94x10^{17}$	-3.25	$1.08x10^{18}$	.00	$1.88x10^{19}$	0.00	.03	1.45
c	2220.0-2279.4	$3.68x10^{21}$	.05	$1.08x10^{21}$	.00	$5.94x10^{17}$	-3.25	$1.08x10^{18}$	.00	$1.88x10^{19}$	0.00	.03	1.45
d	2240.4-2299.8	$3.69x10^{21}$	.04	$1.08x10^{21}$	.00	$5.94x10^{17}$	-3.25	$1.08x10^{18}$	.00	$2.98x10^{19}$	-1.10	.03	1.53

Wg = given concentration in simulated sample, molecules cm^{-2}

Wc = calculated concentration, molecules cm^{-2}

ΔW = Wg-Wc (the factor of a power of ten is omitted), molecules cm^{-2}

See 3.1 for definition of the weighted and unweighted final transmission errors (last two columns). Numbers following gas formulae are computer codes.

A composite sample spectrum was obtained by adding the absorption coefficients as a function of frequency and calculating the transmission (See 1.2). The sample spectrum was then analyzed by MINMYZD to obtain the calculated concentrations. The error is the difference between the calculated and given transmission. (See 3.1).

Table 3.3a
CALCULATION OF GAS CONCENTRATIONS
FROM SIMULATED COMPOSITE SPECTRA (SET 1)
(S/N=10)

Code	Wavenumber Range (cm^{-1})	H_2O (1A1) Wg = 3.73×10^{22} Nc	ΔN	N_2O (4A1) Wg = 1.08×10^{18} Nc	ΔN	CH_4 (5A1) Wg = 1.08×10^{19} Nc	ΔN	SO_2 (20.2) Wg = 3.54×10^{17} Nc	ΔN	HNO_3 (30.2) Wg = 3.54×10^{17} Nc	ΔN	Final Transmission Error RMS Wtd.	Unwtd.
S/N = 10													
8888 a	1230.0-1249.8	3.80×10^{22}	-.07	0.80×10^{18}	.28	1.04×10^{19}	.04	17.88×10^{17}	-14.34	-1.89×10^{17}	5.43	2.08	6.86
b	1250.0-1269.8	3.74×10^{22}	-.01	0.82×10^{18}	.26	1.13×10^{19}	-.05	17.88×10^{17}	-14.34	0.68×10^{17}	2.86	2.28	6.54
c	1270.0-1289.8	3.53×10^{22}	.20	1.04×10^{18}	.04	1.06×10^{19}	.02	17.88×10^{17}	-14.34	4.81×10^{17}	- 1.27	1.86	6.44
d	1290.0-1309.8	3.98×10^{22}	-.25	1.03×10^{18}	.05	1.08×10^{18}	.00	0.28×10^{17}	3.26	3.54×10^{17}	0.00	1.50	1.46
e	1310.0-1329.8	3.98×10^{22}	-.25	0.75×10^{18}	.33	1.02×10^{19}	.06	2.50×10^{17}	1.03	3.08×10^{17}	0.46	0.76	0.54
f	1330.0-1349.8	3.72×10^{22}	.01	1.21×10^{18}	-.13	1.11×10^{19}	-.03	2.75×10^{17}	0.79	3.95×10^{17}	- 0.41	0.80	0.40
g	1350.0-1369.8	3.75×10^{22}	-.02	1.21×10^{18}	-.13	0.99×10^{19}	.09	4.15×10^{17}	- 0.61	4.69×10^{17}	- 1.15	0.28	0.59
h	1370.0-1389.8	3.71×10^{22}	.02	1.21×10^{18}	-.13	1.48×10^{19}	-.40	4.24×10^{17}	- 0.70	18.17×10^{17}	-14.63	0.14	6.55
i	1380.2-1400.0	3.65×10^{22}	.08	1.21×10^{18}	-.13	1.48×10^{19}	-.40	5.60×10^{17}	- 2.06	18.22×10^{17}	-14.68	0.14	6.63

Wg = given concentration in simulated sample, molecules cm^{-2}

Wc = calculated concentration, molecules cm^{-2}

ΔN = Wg-Wc (the factor of a power of ten is omitted), molecules cm^{-2}

See 3.1 for definition of the weighted and unweighted final transmission errors (last two columns). Numbers following gas formulae are computer codes.

A composite sample spectrum was obtained by adding the absorption coefficients as a function of frequency and calculating the transmission (See 1.2). The sample spectrum was then analyzed by MINMYZD to obtain the calculated concentrations. The error is the difference between the calculated and given transmission. (See 3.1.2).

Table 3.3b
(S/N=30)

Code	Wavenumber Range (cm^{-1})	H_2O (1A1) $Wg = 3.73 \times 10^{22}$		N_2O (4A1) $Wg = 1.08 \times 10^{18}$		CH_4 (5A1) $Wg = 1.08 \times 10^{19}$		SO_2 (20.2) $Wg = 3.54 \times 10^{17}$		HNO_3 (30.2) $Wg = 3.54 \times 10^{17}$		Final Transmission Error	
		Wc	ΔW	Wc	ΔW	Wc	ΔW	Wc	ΔW	Wc	ΔW	RMS Wtd.	Unwtd.
S/N = 30													
8888 a	1230.0-1249.8	3.76×10^{22}	-.03	0.50×10^{18}	.58	1.08×10^{19}	.00	17.88×10^{17}	-14.34	0.15×10^{17}	3.39	.65	6.59
b	1250.0-1269.8	3.66×10^{22}	.07	1.23×10^{18}	-.15	1.07×10^{19}	.01	17.88×10^{17}	-14.34	-0.69×10^{17}	4.23	.75	6.69
c	1270.0-1289.8	3.68×10^{22}	.05	1.05×10^{18}	.03	1.05×10^{19}	.03	17.88×10^{17}	-14.34	4.11×10^{17}	- 0.57	.65	6.42
d	1290.0-1309.8	3.73×10^{22}	.00	1.11×10^{18}	-.03	1.07×10^{19}	.01	1.00×10^{17}	2.54	3.50×10^{17}	0.04	.46	1.14
e	1310.0-1329.8	3.75×10^{22}	-.02	1.39×10^{18}	-.30	0.97×10^{19}	.11	4.35×10^{17}	- 0.81	3.40×10^{17}	0.14	.25	0.39
f	1330.0-1349.8	3.73×10^{22}	.00	1.21×10^{18}	-.13	1.05×10^{19}	.03	3.64×10^{17}	- 0.10	3.54×10^{17}	0.00	.26	0.07
g	1350.0-1369.8	3.76×10^{22}	-.03	1.21×10^{18}	-.13	1.07×10^{19}	.01	3.06×10^{17}	0.48	4.58×10^{17}	- 1.04	.10	0.52
h	1370.0-1389.8	3.68×10^{22}	.05	1.21×10^{18}	-.13	1.48×10^{19}	-.40	4.83×10^{17}	- 1.29	18.20×10^{17}	-14.66	.05	6.58
i	1380.2-1400.0	3.69×10^{22}	.04	1.21×10^{18}	-.13	1.48×10^{19}	-.40	4.54×10^{17}	- 1.00	18.24×10^{17}	-14.70	.06	6.59

Wg = given concentration in simulated sample, molecules cm^{-2}

Wc = calculated concentration , molecules cm^{-2}

ΔW = Wg-Wc (the factor of a power of ten is omitted), molecules cm^{-2}

See 3.1 for definition of the weighted and unweighted final transmission errors (last two columns). Numbers following gas formulae are computer codes.

A composite sample spectrum was obtained by adding the absorption coefficients as a function of frequency and calculating the transmission (See 1.2). The sample spectrum was then analyzed by MINMYZD to obtain the calculated concentrations. The error is the difference between the calculated and given transmission. (See 3.1.2).

Table 3.3c
(S/N=100)

Code	Wavenumber Range (cm^{-1})	H_2O (1A1) $Wg = 3.73 \times 10^{22}$		N_2O (4A1) $Wg = 1.08 \times 10^{18}$		CH_4 (5A1) $Wg = 1.08 \times 10^{19}$		SO_2 (20.2) $Wg = 3.54 \times 10^{17}$		HNO_3 (30.2) $Wg = 3.54 \times 10^{17}$		Final Transmission Error	
		Wc	ΔW	Wc	ΔW	Wc	ΔW	Wc	ΔW	Wc	ΔW	RMS Wtd.	Unwtd.
S/N = 100													
8888 a	1230.0-1249.8	$3.74x10^{22}$	-.01	$1.11x10^{18}$	-.03	$1.07x10^{19}$	.01	$17.88x10^{17}$	-14.34	$- 0.03x10^{17}$	3.57	.25	6.61
b	1250.0-1269.8	$3.71x10^{22}$	.02	$1.13x10^{18}$	-.05	$1.07x10^{19}$	.01	$17.88x10^{17}$	-14.34	$0.08x10^{17}$	3.46	.22	6.60
c	1270.0-1289.8	$3.73x10^{22}$	.00	$1.09x10^{18}$	-.01	$1.09x10^{19}$	-.01	$17.88x10^{17}$	-14.34	$3.49x10^{17}$	0.05	.20	6.41
d	1290.0-1309.8	$3.72x10^{22}$	.01	$1.09x10^{18}$	-.01	$1.08x10^{19}$	.00	$0.74x10^{17}$	2.80	$3.50x10^{17}$	0.04	.17	1.25
e	1310.0-1329.8	$3.72x10^{22}$	.01	$1.21x10^{18}$	-.13	$1.08x10^{19}$	.00	$4.11x10^{17}$	- 0.57	$3.50x10^{17}$	0.04	.07	0.26
f	1330.0-1349.8	$3.74x10^{22}$	-.01	$1.21x10^{18}$	-.13	$1.07x10^{19}$	.01	$3.57x10^{17}$	- 0.03	$3.51x10^{17}$	0.03	.09	0.06
g	1350.0-1369.8	$3.74x10^{22}$	-.01	$1.21x10^{18}$	-.13	$1.06x10^{19}$	.02	$3.25x10^{17}$	0.29	$4.37x10^{17}$	- 0.83	.04	0.40
h	1370.0-1389.8	$3.68x10^{22}$	.05	$1.21x10^{18}$	-.13	$1.48x10^{19}$	-.40	$4.84x10^{17}$	- 1.30	$18.37x10^{17}$	-14.83	.03	6.66
i	1380.2-1400.0	$3.67x10^{22}$	.06	$1.21x10^{18}$	-.13	$1.48x10^{19}$	-.40	$5.19x10^{17}$	- 1.65	$18.23x10^{17}$	-14.69	.03	6.61

Wg = given concentration in simulated sample, molecules cm^{-2}

Wc = calculated concentration, molecules cm^{-2}

ΔW = Wg-Wc (the factor of a power of ten is omitted), molecules cm^{-2}

See 3.1 for definition of the weighted and unweighted final transmission errors (last two columns). Numbers following gas formulae are computer codes.

A composite sample spectrum was obtained by adding the absorption coefficients as a function of frequency and calculating the transmission (See 1.2). The sample spectrum was then analyzed by MINMYZD to obtain the calculated concentrations. The error is the difference between the calculated and given transmission. (See 3.1.2).

Table 3.4
CALCULATION OF GAS CONCENTRATIONS FROM SIMULATED COMPOSITE SPECTRA (SET 2)
(S/N = 10, 30, 100)

Code	Wavenumber Range (cm^{-1})	H_2O (1B1) $Wg = 3.73 \times 10^{21}$		CO_2 (2B1) $Wg = 1.08 \times 10^{21}$		O_3 (3B1) $Wg = 2.69 \times 10^{17}$		N_2O (4B1) $Wg = 1.06 \times 10^{18}$		CO (6B1) $Wg = 1.88 \times 10^{19}$		Final Transmission Error	
		Nc	ΔN	Nc	ΔN	Nc	ΔN	Nc	ΔN	Nc	ΔN	RMS Wtd.	Unwtd.
S/N = 10													
9999 a	2100.0-2159.4	3.66×10^{21}	0.07	1.07×10^{21}	.01	1.79×10^{17}	0.90	1.34×10^{18}	-.26	1.91×10^{19}	-.03	1.09	0.42
b	2160.0-2219.4	4.08×10^{21}	-0.35	0.35×10^{21}	.73	5.94×10^{17}	-3.25	1.07×10^{18}	.01	1.88×10^{19}	.00	1.28	1.50
c	2220.0-2279.4	5.10×10^{21}	-1.37	1.11×10^{21}	-.03	5.94×10^{17}	-3.25	1.11×10^{18}	-.05	1.87×10^{19}	.01	2.64	1.58
d	2240.4-2299.8	-0.93×10^{21}	4.66	1.10×10^{21}	-.02	5.94×10^{17}	-3.25	1.06×10^{18}	.02	1.33×10^{19}	.55	3.54	2.55
S/N = 30													
9999 a	2100.0-2159.4	3.73×10^{21}	0.00	1.14×10^{21}	-.06	2.53×10^{17}	0.16	1.19×10^{18}	-.11	1.88×10^{19}	.00	0.36	0.09
b	2160.0-2219.4	3.89×10^{21}	-0.16	1.27×10^{21}	-.19	5.94×10^{17}	-3.25	1.08×10^{18}	.00	1.88×10^{19}	.00	0.44	1.46
c	2220.0-2279.4	4.72×10^{21}	-0.99	1.09×10^{21}	-.01	5.94×10^{17}	-3.25	1.09×10^{18}	-.01	1.89×10^{19}	-.01	0.86	1.52
d	2240.4-2299.8	2.67×10^{21}	1.06	1.08×10^{21}	.00	5.94×10^{17}	-3.25	1.07×10^{18}	.01	1.49×10^{19}	.39	1.19	1.54
S/N = 100													
9999 a	2100.0-2159.4	3.72×10^{21}	0.01	1.07×10^{21}	.01	3.01×10^{17}	-.32	1.19×10^{18}	-.11	1.88×10^{19}	.00	0.10	0.15
b	2160.0-2219.4	3.68×10^{21}	0.05	1.08×10^{21}	.00	5.94×10^{17}	-3.25	1.08×10^{18}	.00	1.88×10^{19}	.00	0.14	1.45
c	2220.0-2279.4	3.68×10^{21}	0.05	1.08×10^{21}	.00	5.94×10^{17}	-3.25	1.08×10^{18}	.00	1.87×10^{19}	.01	0.26	1.45
d	2240.4-2299.8	3.78×10^{21}	-0.05	1.08×10^{21}	.00	5.94×10^{17}	-3.25	1.08×10^{18}	.00	2.59×10^{19}	-.71	0.31	1.49

Wg = given concentration in simulated sample, molecules cm^{-2}

Nc = calculated concentration, molecules cm^{-2}

ΔN = Wg-Nc (the factor of a power of ten is omitted), molecules cm^{-2}

See 3.1 for definition of the weighted and unweighted final transmission errors (last two columns). Numbers following gas formulae are computer codes.

A composite sample spectrum was obtained by adding the absorption coefficients as a function of frequency and calculating the transmission (See 1.2). The sample spectrum was then analyzed by MINMY2D to obtain the calculated concentrations. The error is the difference between the calculated and given transmission. (See 3.1.2).

To apply directly the method to a sample of unknown composition, utilizing the full library of test spectra, would be an example of how much must be paid (in computer costs) for total ignorance of the sample. Maximum information available on the composition of the unknown sample should be utilized in selecting the library of test spectra. In the examples discussed below (3.1.3), we will see the penalty on the accuracy of the calculation for omitting a spectrum from the library when the gas is in the sample.

Further, we note that suitable choices of frequency regions to investigate the presence of a given gas must be made. Obviously, if at concentration levels of interest, a gas shows no (or weak) absorption, its concentration may not be measured there. In F3.1 (and F1.1a) consider the two spectral regions of absorption by HNO_3, namely those at 880 and 1300 cm^{-1}. The higher frequency band has strong potential interference from H_2O, SO_2, N_2O, CH_4, and all HC. Thus, if HNO_3 is suspected in a sample, the concentration should be calculated in the lower frequency band. Therefore, the analyst should consult the positioning of spectra in F3.1 before beginning his analysis.

Suppose an analyst had unlimited free computer time. Should he not then simplify the exercise of judgment and insert his entire library into the computer and analyze the whole frequency range? But here he will find a requirement for careful judgment also. For, as we will see below, in regions where absorption is effectively zero, the results of the calculation are small but essentially random. Therefore, the analyst must choose from among the calculated concentrations obtained in all regions only those where the absorption is substantial for the gas of interest. Such a weighting of values could be added to the computer program. It is uncertain whether the advantages outweigh the added complexity in the program.

In T3.1 through T3.3, a uniform notation is used. Two separate sets of gases are used. Set 1 consists of H_2O, N_2O, CH_4, SO_2, and HNO_3, and has the Code 8888. Set 2 consists of H_2O, CO_2, O_3, N_2O, and CO,

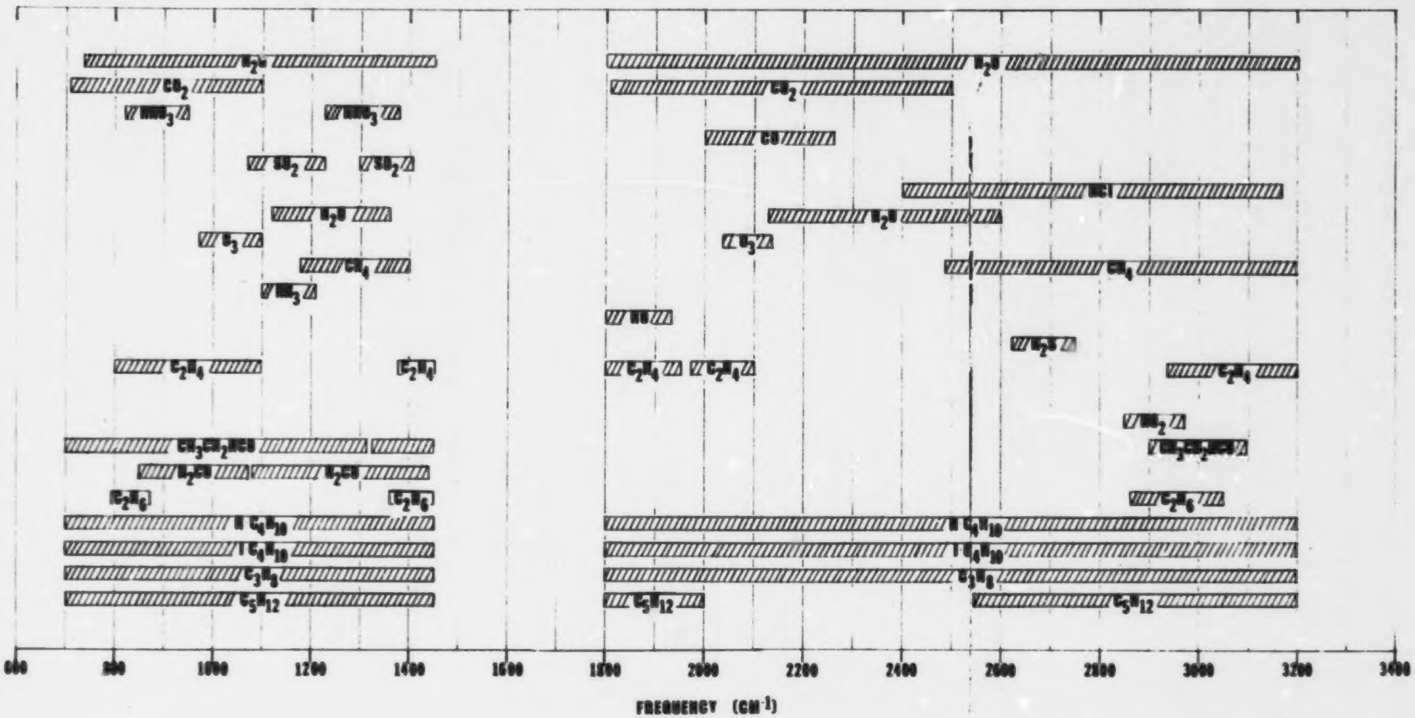


Figure 3.1

PRINCIPAL REGIONS OF ABSORPTION OF
SELECTED NORMAL AND POLLUTANT ATMOSPHERIC CONSTITUENTS
(The region from 1460 to 1800 cm^{-1} has
strong water vapor absorption)

and has the Code 9999. H_2O is included in both sets because it is always present in the atmosphere; CO_2, though always present, is not as variable. The other gases were chosen to be representative spectra in the upper and lower frequency ranges (see F1 (a-c) and F3.1). The concentration (molecules cm^{-2}) is denoted for each gas by Wg. The concentrations are chosen to match representative values in the Los Angeles area. The frequency range for which the calculation was made is designated by a small letter. The total range was divided into subranges to observe the effect of spectral strength and structure on the accuracy of the computation. The calculated concentration is denoted by Wc. The difference $(Wg-Wc)x10^{-y}$ is a measure of the accuracy (y is the exponent in Wg). Two measures of rms error are given, the weighted error and the unweighted error. The weighted error is of use to the programmer. Here we will refer to the unweighted error.

The penultimate columns of the tables contain those final transmission errors defined as

$$\left[\frac{1}{N}\sum_{k=1}^{N}\left\{\left[(\text{True Transmission})_k-(\text{Calculated Transmission})_k\right]CF_k\right\}^{2}\right]^{1/2},$$

where CF_k is a "confidence factor" introduced to weight in favor of high concentrations because they produce, generally speaking, less error. (See A2.5.) The transmissions involved are those due to the total sample of all gases. The summation is over each experimental data value of the frequency range given in the tables. It thus is a figure of merit for "goodness of fit" over that range. The unweighted error is obtained by letting all $CF_k=1$.

The unweighted error presented in the final column is a different sort of measure. It is simply the rms of the fractional errors (Wg-Wc)/Wg, on a line of the table. In the error column, factors of 10 have been omitted.

Tables 3.1 and 3.2 are based on an assumption of infinite signal-to-noise ratio. Tables 3.3 (a-c) and 3.4 illustrate the effect of finite S/N. Tables 3.5 (a, b) and 3.6 illustrate the effect of "maverick" gases. All these cases are discussed in the following sections.

3.1.1 Calculation of Concentrations for Spectra Without Noise

Consider the calculations coded as 8888 (a-i) in T3.1 (Set 1). A simulated sample was prepared, containing concentrations of the gases listed, unknown to the analyst. Using the simulated sample as data, concentrations were computed by the MINMYZD method. In the frequency ranges given in the second column, a concentration of water of $Wg = 3.73 \times 10^{22}$ is relatively strong and the small errors (ΔW) under H_2O are to be expected when, as here, the S/N is assumed infinite. The inaccuracy lies between 0 and 2 percent, the larger error occurring toward the strong "black-out" absorption band of H_2O lying between 1430 and 1800 cm^{-1}. Interference from the other constituents of the sample has little effect on the calculated concentrations of H_2O.

The calculated concentration of N_2O from the data is done with an error not exceeding 14 percent. The larger error occurs where the absorption is very small, and such regions should be lightly weighted in a final averaging. By such weighting, we may estimate a maximum error of 5 percent in this case.

Methane (CH_4) has a strong absorption spectrum and very distinctive spectral structure in part of the frequency range covered in T3.1, and the error in the calculated concentration is small except where the absorption is very weak in the range 1370 to 1400 cm^{-1}, being 37 percent.

The given concentration of SO_2, $Wg = 3.54 \times 10^{17}$ molecules cm^{-2}, is relatively small. Better results than shown can be obtained by using the stronger absorption spectrum between 1050 and 1200 cm^{-1} (see F1.1a). Here, the errors, excluding the regions from 1230 to 1310 cm^{-1}, range from 1 to 44 percent. By using the results only from the strongest absorbing region, the errors are quite small, averaging 9 percent.

The last gas considered in this set is HNO_3. Its spectrum is strong, but the concentration is relatively low. Again we see small errors where the absorption is strong, and large errors amounting to 100 percent where the absorption is weak.

The calculations listed under Codes 9999 (a-d) in T3.2 give a similar set of calculations for a different set of gases, namely H_2O, CO_2, O_3, N_2O, and CO. The spectra of these gases are shown in F1.1a, except for O_3. Its absorption, in the frequency range 2100 to 2300 cm^{-1} , is too weak to show on the scale of the graph; it is listed however in the library of spectra.

In this set of gases, the errors, except for O_3, are quite small. The best analysis of O_3 would be made in the stronger absorption band centered at 1100 cm^{-1}.

It is, of course, not surprising that we may conclude from these computations that errors for a particular gas are minimized by choosing regions of strong absorption of that gas, as free from interference as possible. These results illustrate the earlier discussion that intelligent choices of frequency regions for the computation are necessary.

3.1.2 Effect of Finite Signal-to-Noise Ratio (S/N)

As remarked in section 1.1, the spectrophotometer system has a finite S/N ratio. To estimate the effect of noise on the system, simulated random noise has been added to the data in runs under Codes 8888 (a-i) and 9999 (a-d) as shown in T3.3 (a-c) and T3.4. The S/N ratios of 10, 30, and 100 were simulated. The method of simulation is discussed in A2.

By comparing T3.1 and T3.3, we see that errors (ΔW) are not substantially increased at SN = 100 over the noise-free case. Average errors increase at S/N = 30, and they become considerable in some cases at S/N = 10. Generally, the calculated concentrations are within an order of magnitude of the true sample concentrations (except in the case of SO_2 and HNO_3 in those regions of the spectrum where absorption is weak). Whether these errors are tolerable or not depends on the needs of the experimenter. In many cases, order of magnitude results are all that are needed for preliminary assessments of a polluted atmosphere. Greater accuracy could be achieved by longer paths and longer integration times, provided that the instrument is stable against long period drift and the sample is stationary.

Even at S/N = 10, the errors may be kept low by choosing regions where the spectrum is strong. It is certainly true, however, that as S/N becomes much less than 10, the accuracy will decrease significantly.

Table 3.4 may be compared with T3.2 for the effect of noise on concentrations calculated in the spectral range from 2100-2300 cm^{-1}. The gases (Set 2) H_2O, CO_2, O_3, N_2O, and CO are present in the simulated sample (Codes 9999 (a-d)). Here conclusions parallel those of the above discussion; a signal-to-noise ratio more than 10 is required to keep errors below 20 percent, and S/N = 100 produces negligible errors.

The reader should not be perturbed because a few low S/N cases appear more accurate than those with higher or infinite S/N. This occurs for two reasons: (1) the noise added may accidently provide an occasional better fit; and (2) the repetition of the algorithm seeking the best fit is at the discretion of the operator. In the case of low S/N, the search was allowed to proceed for more iterations than in the less noisy cases in order to provide the best opportunity of overcoming the noise.

3.1.3 Effects of "Maverick" Gases in the Sample

In analysis of a spectrum from a specific geographical location, it is important to have in the library of spectra, utilized in the computer analysis, a spectrum of each of the gases expected in the area. However, the library should be kept as small as possible to minimize computer storage problems. Thus, the library data entered into the computer memory are chosen to match as closely as possible the set of gases that are in the atmosphere under measurement. For a larger than necessary library (i.e., some gases in the library are not in the test atmosphere), the only important penalties are increased computer memory required and the additional costs due to the necessity of breaking the frequency spectrum into separately treated sections, each run separately on the computer, if the limited computer memory requires it.

To test the effect of having gases present in the sample, but not present in the computer library (a "maverick" gas), runs were made on both previously used sets of gases. From each set of calculations, spanning for Set 1, from 1230 to 1400 cm^{-1} in nine steps, each gas, except H_2O, was successively deleted from the library (T3.5a,b). Similar procedures were followed in the 2100 to 2300 cm^{-1} range (T3.6).

The gas deleted from the library was included in the simulated data, but was not among those which the computer attempted to fit to the sample data by best choice of concentration. It thus acted as a localized noise or perturbation of the spectrum.

Comparing T3.1 with T3.5 (a,b) and T3.2 with T3.6, we see that appreciable errors may arise by having an insufficient library. The greatest of these appear where the spectrum of the maverick gas overlaps a weak spectrum of a gas whose spectrum is in the library. This result indicates that as a practical measure it is better to add a gas to the library if its presence in the sample is uncertain.

An examination of the spectra of Sets 1 and 2 in F1.1 (a-c) and F3.1 shows that the spectral structure must act as an appreciable noise background where the gas is a maverick. The MINMYZD method, by utilizing goodness of fit over a wide spectral region, partially eliminates this noise effect.

When a gas spectrum is present in the library, but near or at zero concentration in the sample, the calculation occasionally produces negative concentrations, small in absolute value. They could be prevented by a simple change in the program, but it was judged preferable not to affect the convergence process by putting a barrier at zero concentrations. Such negative concentrations should be considered equal to zero. This situation is discussed in A2.5.

Table 3.5a

GAS CONCENTRATIONS FROM SIMULATED COMPOSITE SPECTRA (SET 1)

Code	Wavenumber Range (cm^{-1})	H_2O (1A1) Wg = 3.73×10^{22} Wc	ΔW	N_2O (4A1) Wg = 1.08×10^{18} Wc	ΔW	CH_4 (5A1) Wg = 1.08×10^{19} Wc	ΔW	SO_2 (20.2) Wg = 3.54×10^{17} Wc	ΔW	HNO_3 (30.2) Wg = 3.54×10^{17} Wc	ΔW	Final Transmission Error RMS Wtd.	Unwtd.
8888 a,B	1230.0-1249.8	3.78×10^{22}	-0.05			1.10×10^{19}	-.02	17.88×10^{17}	-14.34	-0.07×10^{17}	3.61	0.35	7.39
b,B	1250.0-1269.8	4.00×10^{22}	-0.27			1.18×10^{19}	-.10	17.88×10^{17}	-14.34	24.58×10^{17}	-21.04	0.74	12.73
c,B	1270.0-1289.8	3.92×10^{22}	-0.19	N_2O deleted from library for runs 8888a,B through 8888i,B		1.02×10^{19}	.06	17.88×10^{17}	-14.34	8.10×10^{17}	- 4.56	1.76	7.52
d,B	1290.0-1309.8	4.04×10^{22}	-0.31			1.34×10^{19}	-.26	-0.18×10^{17}	3.72	5.81×10^{17}	- 2.27	0.60	2.19
e,B	1310.0-1329.8	3.69×10^{22}	0.04			1.31×10^{19}	-.23	0.42×10^{17}	3.12	4.85×10^{17}	- 1.31	0.11	1.70
f,B	1330.0-1349.8	3.73×10^{22}	0.00			1.08×10^{19}	.00	3.51×10^{17}	0.03	3.56×10^{17}	- 0.02	0.02	0.02
g,B	1350.0-1369.8	3.73×10^{22}	0.00			1.06×10^{19}	.02	3.32×10^{17}	0.22	4.32×10^{17}	- 0.78	0.03	0.41
h,B	1370.0-1389.8	3.68×10^{22}	0.05			1.48×10^{19}	-.40	4.88×10^{17}	- 1.34	18.38×10^{17}	-14.84	0.03	7.48
i,B	1380.2-1400.0	3.67×10^{22}	0.06			1.48×10^{19}	-.40	5.10×10^{17}	- 1.56	18.23×10^{17}	-14.69	0.03	7.39
8888 a,B1	1230.0-1249.8	4.66×10^{22}	-0.93	8.51×10^{18}	-7.43			17.88×10^{17}	-14.34	25.31×10^{17}	-21.77	9.19	18.[illegible]
b,B1	1250.0-1269.8	3.39×10^{22}	0.34	4.85×10^{18}	-3.77			17.88×10^{17}	-14.34	-17.90×10^{17}	21.44	7.42	13.[illegible]
c,B1	1270.0-1289.8	2.90×10^{22}	0.83	-0.59×10^{18}	1.67	CH_4 deleted from library for runs 8888a,B1 through 8888i,B1		17.88×10^{17}	-14.34	25.05×10^{17}	-21.51	7.79	12.[illegible]
d,B1	1290.0-1309.8	2.06×10^{22}	1.67	5.52×10^{18}	-4.44			-1.12×10^{17}	4.66	-2.27×10^{17}	5.81	10.36	4.48
e,B1	1310.0-1329.8	4.56×10^{22}	-0.83	3.71×10^{18}	-2.63			10.34×10^{17}	- 6.80	-0.27×10^{17}	3.81	0.72	4.18
f,B1	1330.0-1349.8	4.09×10^{22}	-0.36	1.21×10^{18}	-0.13			0.32×10^{17}	3.22	9.18×10^{17}	- 5.64	0.71	3.29
g,B1	1350.0-1369.8	4.33×10^{22}	-0.60	1.21×10^{18}	-0.13			0.96×10^{17}	2.58	24.01×10^{17}	-20.47	0.74	10.32
h,B1	1370.0-1389.8	3.68×10^{22}	0.05	1.21×10^{18}	-0.13			4.88×10^{17}	- 1.34	18.38×10^{17}	-14.84	0.03	7.48
i,B1	1380.2-1400.0	3.67×10^{22}	0.06	1.21×10^{18}	-0.13			5.10×10^{17}	- 1.56	18.23×10^{17}	-14.69	0.03	7.39

Wg = given concentration in simulated sample, molecules cm^{-2}

Wc = calculated concentration, molecules cm^{-2}

ΔW = Wg-Wc (the factor of a power of ten is omitted), molecules cm^{-2}

See 3.1 for definition of the weighted and unweighted final transmission errors (last two columns). Numbers following gas formulae are computer codes.

A composite sample spectrum was obtained by adding the absorption coefficients as a function of frequency and calculating the transmission (See 1.2). The sample spectrum was then analyzed by MINMYZD to obtain the calculated concentrations. The error is the difference between the calculated and given transmission. (See 3.1.3).

Table 3.5b

Code	Wavenumber Range (cm^{-1})	H_2O (1A1) $Wg = 3.73 \times 10^{22}$		N_2O (4A1) $Wg = 1.08 \times 10^{18}$		CH_4 (5A1) $Wg = 1.08 \times 10^{19}$		SO_2 (20.2) $Wg = 3.54 \times 10^{17}$		HNO_3 (30.2) $Wg = 3.54 \times 10^{17}$		Final Transmission Error	
		Wc	ΔW	Wc	ΔW	Wc	ΔW	Wc	ΔW	Wc	ΔW	RMS Wtd.	Unwtd.
8888 a,B2	1230.0-1249.8	3.73×10^{22}	.00	1.09×10^{18}	-0.01	1.08×10^{19}	.00			0.07×10^{17}	3.47	0.03	1.74
b,B2	1250.0-1269.8	3.73×10^{22}	.00	1.09×10^{18}	-0.01	1.08×10^{19}	.00			0.10×10^{17}	3.44	0.06	1.72
c,B2	1270.0-1289.8	3.73×10^{22}	.00	1.08×10^{18}	0.00	1.08×10^{19}	.00			3.55×10^{17}	- 0.01	0.03	0.01
d,B2	1290.0-1309.8	3.73×10^{22}	.00	1.08×10^{18}	0.00	1.08×10^{19}	.00	SO_2 deleted from		3.54×10^{17}	0.00	0.03	0.00
e,B2	1310.0-1329.8	3.64×10^{22}	.09	0.31×10^{18}	0.77	1.16×10^{19}	-.08	library for runs		5.19×10^{17}	- 1.65	0.11	0.91
f,B2	1330.0-1349.8	4.05×10^{22}	-.32	1.21×10^{18}	-0.13	1.06×10^{19}	.02	8888a,B2, through		4.50×10^{17}	- 0.96	0.36	0.51
g,B2	1350.0-1369.8	4.08×10^{22}	-.35	1.21×10^{18}	-0.13	1.09×10^{19}	-.01	8888i,B2		7.82×10^{17}	- 4.28	0.06	2.15
h,B2	1370.0-1389.8	4.04×10^{22}	-.31	1.21×10^{18}	-0.13	1.48×10^{19}	-.40			18.25×10^{17}	-14.71	0.02	7.36
i,B2	1380.2-1400.0	4.04×10^{22}	-.31	1.21×10^{18}	-0.13	1.48×10^{19}	-.40			18.23×10^{17}	-14.69	0.02	7.35
8888 a,B3	1230.0-1249.8	3.73×10^{22}	.00	1.09×10^{18}	-0.01	1.08×10^{19}	.00	17.88×10^{17}	-14.34			0.03	7.17
b,B3	1250.0-1269.8	3.73×10^{22}	.00	1.09×10^{18}	-0.01	1.09×10^{19}	.00	17.88×10^{17}	-14.34			0.06	7.17
c,B3	1270.0-1289.8	4.05×10^{22}	-.32	1.32×10^{18}	-0.24	1.22×10^{19}	-.14	17.88×10^{17}	-14.34	HNO_3 deleted from		0.72	7.17
d,B3	1290.0-1309.8	3.53×10^{22}	.20	3.02×10^{18}	-1.94	0.86×10^{19}	.22	14.24×10^{17}	-10.70	library for runs		1.76	5.44
e,B3	1310.0-1329.8	4.34×10^{22}	-.61	2.93×10^{18}	-1.85	0.82×10^{19}	.26	11.80×10^{17}	- 8.26	8888a,B3 through		0.19	4.25
f,B3	1330.0-1349.8	4.00×10^{22}	-.27	1.21×10^{18}	-0.13	1.42×10^{19}	-.34	10.10×10^{17}	- 6.56	8888i,B3		0.81	3.29
g,B3	1350.0-1369.8	3.64×10^{22}	.09	1.21×10^{18}	-0.13	1.07×10^{19}	.01	6.14×10^{17}	- 2.60			0.04	1.30
h,B3	1370.0-1389.8	3.68×10^{22}	.05	1.21×10^{18}	- .13	1.48×10^{19}	-.40	4.86×10^{17}	- 1.32			0.03	0.69
i,B3	1380.2-1400.0	3.67×10^{22}	.06	1.21×10^{18}	- .13	1.48×10^{19}	-.40	5.10×10^{17}	- 1.56			0.03	0.81

Wg = given concentration in simulated sample, molecules cm^{-2}

Wc = calculated concentration, molecules cm^{-2}

ΔW = Wg-Wc (the factor of a power of ten is omitted), molecules cm^{-2}

See 3.1 for definition of the weighted and unweighted final transmission errors (last two columns). Numbers following gas formulae are computer codes.

A composite sample spectrum was obtained by adding the absorption coefficients as a function of frequency and calculating the transmission (See 1.2). The sample spectrum was then analyzed by MINMYZD to obtain the calculated concentrations. The error is the difference between the calculated and given transmission. (See 3.1.3).

Table 3.6
GAS CONCENTRATIONS FROM SIMULATED COMPOSITE SPECTRA (SET 2)
(MAVERICK GASES)

Code	Wavenumber Range (cm^{-1})	H_2O (1B1) Wg = 3.73×10^{21}		CO_2 (2B1) Wg = 1.08×10^{21}		O_3 (3B1) Wg = 2.69×10^{17}		N_2O (4B1) Wg = 1.08×10^{18}		CO (6B1) Wg = 1.88×10^{19}		Final Transmission Error	
		Wc	ΔW	Wc	ΔW	Wc	ΔW	Wc	ΔW	Wc	ΔW	RMS Wtd.	Unwtd.
9999 a,A	2100.0-2159.4	3.65×10^{21}	0.08	CO_2 deleted from		5.07×10^{17}	- 2.38	0.64×10^{18}	0.44	1.93×10^{19}	- 0.05	0.43	1.21
b,A	2160.0-2219.4	3.73×10^{21}	0.00	library for runs		5.94×10^{17}	- 3.25	1.08×10^{18}	0.00	1.88×10^{19}	0.00	0.03	1.63
c,A	2220.0-2279.4	49.64×10^{21}	-45.91	9999a,A through		5.94×10^{17}	- 3.25	0.93×10^{18}	0.15	1.78×10^{19}	0.10	24.22	23.01
d,A	2240.4-2299.8	11.74×10^{21}	- 8.01	9999d,A		4.40×10^{17}	- 1.71	0.51×10^{18}	0.57	2.98×10^{19}	- 1.10	63.50	4.14
9999 a,A1	2100.0-2159.4	3.68×10^{21}	0.05	1.29×10^{21}	- 0.21	O_3 deleted from		0.81×10^{18}	0.27	1.91×10^{19}	- 0.03	0.20	0.17
b,A1	2160.0-2219.4	3.72×10^{21}	0.01	1.05×10^{21}	0.03	library for runs		1.08×10^{18}	0.00	1.88×10^{19}	0.00	0.03	0.02
c,A1	2220.0-2279.4	3.68×10^{21}	0.05	1.08×10^{21}	0.00	9999a,A1 through		1.08×10^{18}	0.00	1.88×10^{19}	0.00	0.03	0.03
d,A1	2240.4-2299.8	3.69×10^{21}	0.04	1.08×10^{21}	0.00	9999d,A1		1.08×10^{18}	0.00	2.98×10^{19}	- 1.10	0.03	0.55
9999 a,A2	2100.0-2159.4	3.72×10^{21}	0.01	1.05×10^{21}	0.03	2.64×10^{17}	0.05	N_2O deleted from		1.89×10^{19}	- 0.01	0.04	0.03
b,A1	2160.0-2219.4	1.58×10^{21}	2.15	5.76×10^{12}	- 4.68	5.94×10^{17}	- 3.25	library runs		3.88×10^{19}	- 2.00	24.13	3.20
c,A2	2220.0-2279.4	81.24×10^{21}	-77.51	0.80×10^{21}	0.28	5.94×10^{17}	- 3.25	9999a,A2 through		14.11×10^{19}	-12.23	33.67	39.27
d,A2	2240.4-2299.8	11.74×10^{21}	- 8.01	0.92×10^{21}	0.16	4.40×10^{17}	- 1.71	9999d,A2		2.98×10^{19}	- 1.10	29.44	4.13
9999 a,A3	2100.0-2159.4	12.55×10^{21}	- 8.82	11.11×10^{21}	-10.03	45.48×10^{17}	-42.79	44.14×10^{18}	-43.06	CO deleted from		16.55	31.08
b,A3	2160.0-2219.4	36.19×10^{21}	-32.46	-5.34×10^{21}	6.42	5.94×10^{17}	- 3.25	1.79×10^{18}	- 0.71	library for runs		18.60	16.63
c,A3	2220.0-2279.4	3.59×10^{21}	0.14	1.08×10^{21}	0.00	5.94×10^{17}	- 3.25	1.14×10^{18}	- 0.06	9999a,A3 through		2.06	1.63
d,A3	2240.4-2299.8	3.61×10^{21}	0.12	1.08×10^{21}	0.00	5.94×10^{17}	- 3.25	1.08×10^{18}	0.00	9999d,A3		0.03	1.63

Wg = given concentration in simulated sample, molecules cm^{-2}

Wc = calculated concentration, molecules cm^{-2}

ΔW = Wg-Wc (the factor of a power of ten is omitted), molecules cm^{-2}

See 3.1 for definition of the weighted and unweighted final transmission errors (last two columns). Numbers following gas formulae are computer codes.

A composite sample spectrum was obtained by adding the absorption coefficients as a function of frequency and calculating the transmission (See 1.2). The sample spectrum was then analyzed by MINMYZD to obtain the calculated concentrations. The error is the difference between the calculated and given transmission. (See 3.1.3).

4. CALCULATION OF THE SELECTED GAS SAMPLE CONCENTRATIONS FROM SPECTROPHOTOMETER DATA

The results presented in 3 were obtained from simulated data samples produced by calculating composite spectra from selected gas spectra for specified concentrations. While these are an adequate test of the computer program, they do not test the overall system of spectrophotometer and computer algorithm. Accordingly, composite spectra were prepared by Dr. Wm. Herget, EPA, Research Triangle Park, NC, on the ROSE* system described in 1.1 and A3. These data were recorded on digital magnetic tape, and this information became the input to the EPAGAS program described in A2. In addition, for comparison, some data were obtained on a Fourier transform spectrometer (FTS).

The data were not obtained from a long atmospheric path because of the difficulty of obtaining, in that case, an independent measure of the gaseous concentrations. Instead, a cell containing approximately measured concentrations of selected gases was placed in the path. The data were normalized, approximately, by removing the cell and running over the same spectral range to determine the spectra of ambient H_2O and CO_2.

Several problems arose in connection with normalization and wavelength calibration. These are discussed in 1.1 and A2.5, respectively. The result of normalization and wavelength uncertainties in the spectrophotometer output and the added effect of mismatched resolutions of the spectrophotometer and the spectra library result in some small inaccuracies in the calculated concentrations. These problems can all be solved by better calibration methods, and effort is underway to do so. The ROSE spectrophotometer resolution is a function of wavelength, which is characteristic of grating spectrometers. It can be prevented by a nonlinear slit drive mechanism or by using a Fourier Transform Spectrometer (FTS). The FTS can be expected to have a linear frequency scale, but the data may need a shift of frequency to correct for an uncertain zero calibration. The ROSE spectrophotometer data and the data from the FTS correspond well in general. However, in a number of regions, the correspondence is sufficiently poor to cause errors in calculated concentrations. Data with such discrepancies may be seen by comparing F4.1, F4.2, and F4.3 with F4.9 (a-c).

---

*Remote Optical Sensing of Emissions.

However, in spite of these difficulties, the computational method performs well, and as shown below, the results of the calculation are of good accuracy when precautions are taken to calibrate the data.

Several calculations from experimental data are described in the following sections, utilizing different gases in different frequency ranges.

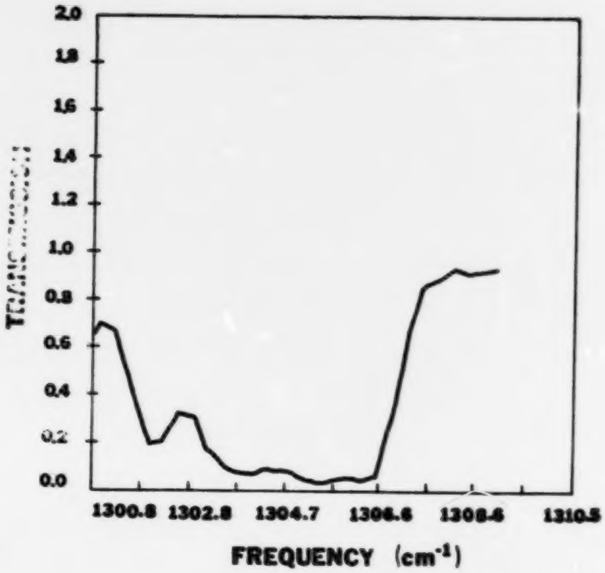


Figure 4.1

SPECTRUM OF SO_2 AND CH_4 OBTAINED BY FTS (1300.8-1310.5 cm^{-1})

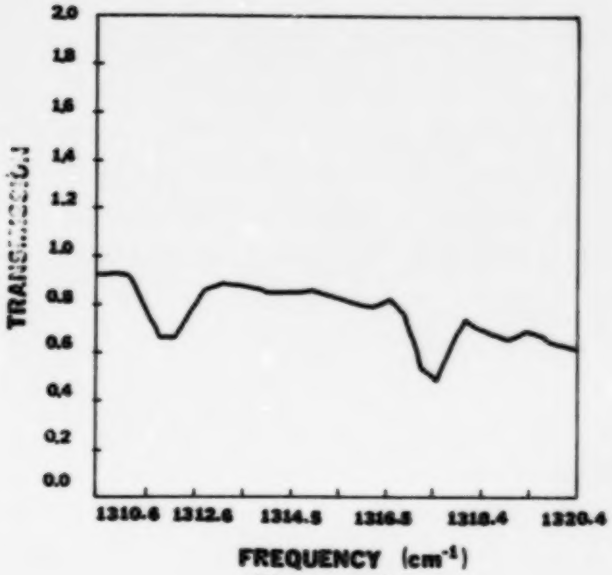


Figure 4.2

SPECTRUM OF SO_2 AND CH_4 OBTAINED BY FTS (1310.6-1320.4 cm^{-1})

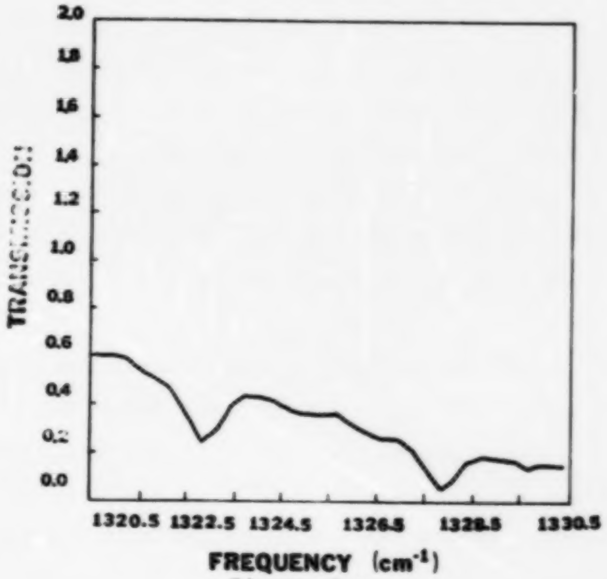


Figure 4.3

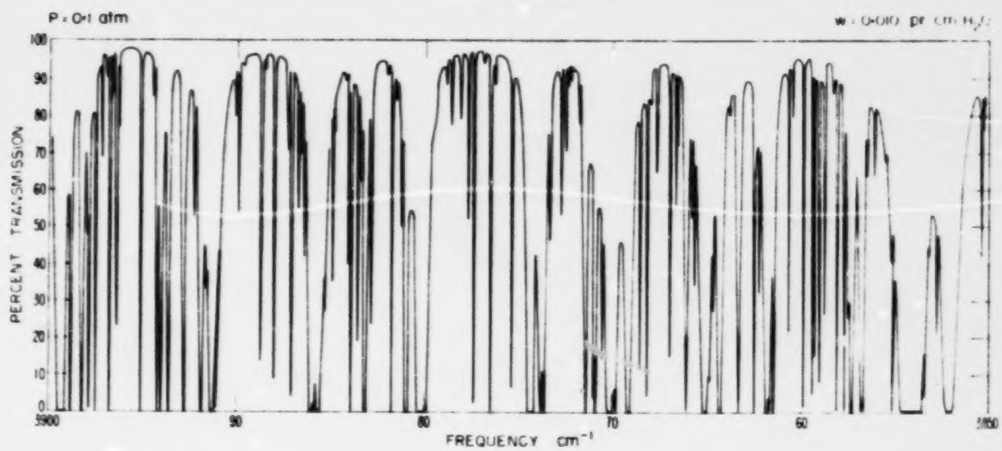
SPECTRUM OF SO_2 AND CH_4 OBTAINED BY FTS (1320.5-1330.5 cm^{-1})

4.1 EFFECT OF MISMATCHED RESOLUTION OF EXPERIMENTAL DATA AND LIBRARY SPECTRA

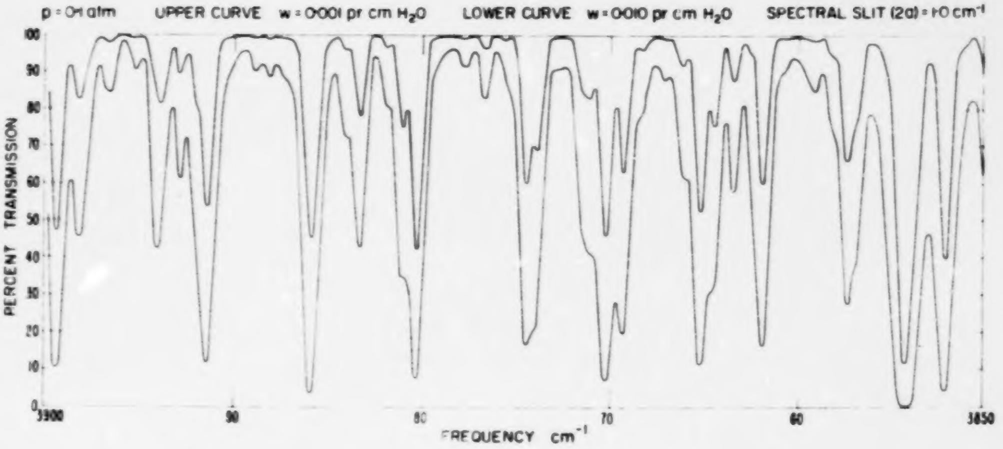
The resolution of spectroscopic instruments has a large effect on the appearance of molecular spectra. It is well known that instrumental resolution can change the shape of a theoretical spectrum so thoroughly that recognition is very difficult. Appearance of the spectrum is of great importance in any best-fit program because it compares two curves, the one obtained from the spectral library with the one from the data. Figure 4.4 shows an example of a theoretical spectrum degraded by an instrumental bandwidth (Calfee, 1966).

The gross changes of recognizable features seen in F4.4 are not usually a problem in spectral analysis of pollutants because the mismatch in resolution is not as large. However, considerable error in analysis can be introduced by such mismatch. Figure 4.5 is a position of the reconstructed spectrum of carbon monoxide reproduced from data obtained on the ROSE system. The concentration of the sample was 1.5×10^{19} molecules cm^{-2} and the spectral slit width was 3 cm^{-1} according to the accompanying data. Difficulty was encountered in analyzing the spectrum, hence a series of computed spectra were made with various amounts of gas and with different slit widths. Figure 4.6 was computed using a spectral slit width of 3 cm^{-1} and with 1.5×10^{18}, 7.5×10^{18}, and 1.5×10^{19} molecules cm^{-2} to see if a different amount of gas in the cell could be the cause of the discrepancy. This was not the case because none of the computed spectra match the data.

Next, a calculation was made with the same quantities of gas, but with a narrower slit width (2.75 cm^{-1}). Figure 4.7 shows the result of this calculation. This result approaches closer to the experimental spectrum. Figure 4.8 was made by using 2.3 cm^{-1} as the spectral slit width. Here the agreement between the computed spectrum and the experimental curve is very good for a concentration of 1.5×10^{19} molecules cm^{-2}. This analysis points up the significance of the spectral slit width in the structure of the spectrum and the best-fit procedure. (The slit width as a function of frequency is shown in A3.)



THEORETICAL SPECTRUM



DEGRADED SPECTRUM

Figure 4.4

COMPARISON OF A THEORETICAL SPECTRUM (HIGH RESOLUTION) OF WATER VAPOR AND THE SPECTRUM OBSERVED THROUGH A SPECTROMETER WITH 1.0 cm^{-1} SLIT WIDTH (IN MANY REGIONS THE COMPARISON IS VERY DIFFICULT)

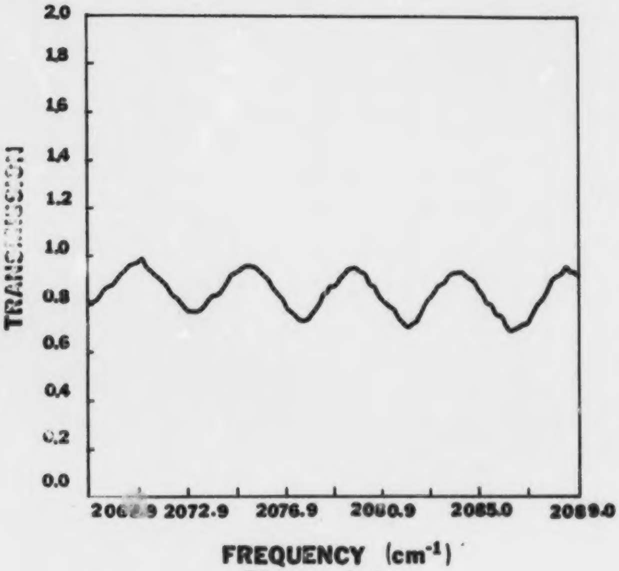


Figure 4.5

EXPERIMENTAL SPECTRUM, CO,
NOMINAL RESOLUTION 2.5 cm^{-1}
OBTAINED ON ROSE SYSTEM
(Compare with F4.6)

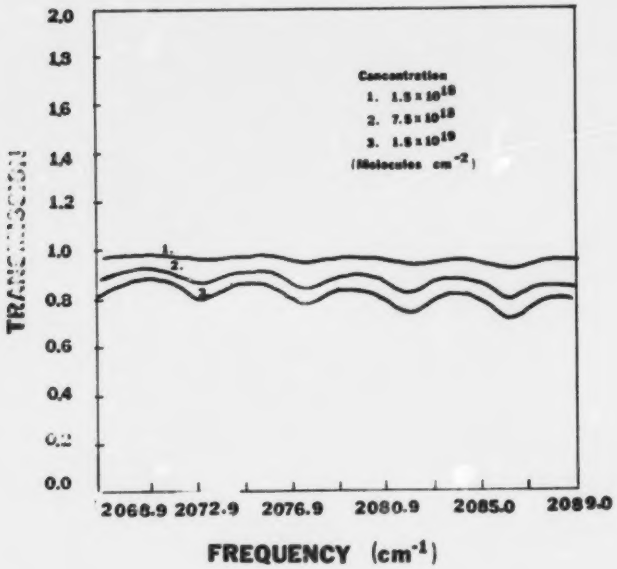


Figure 4.6

CALCULATED SPECTRUM, CO,
RESOLUTION 3.0 cm^{-1}
(Compare with F4.5)

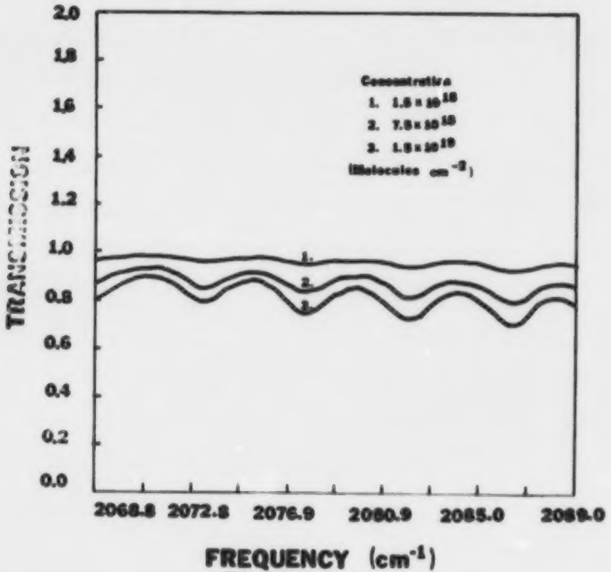


Figure 4.7

CALCULATED SPECTRUM, CO,
RESOLUTION 1.0 cm^{-1}
(Compare with F4.5)

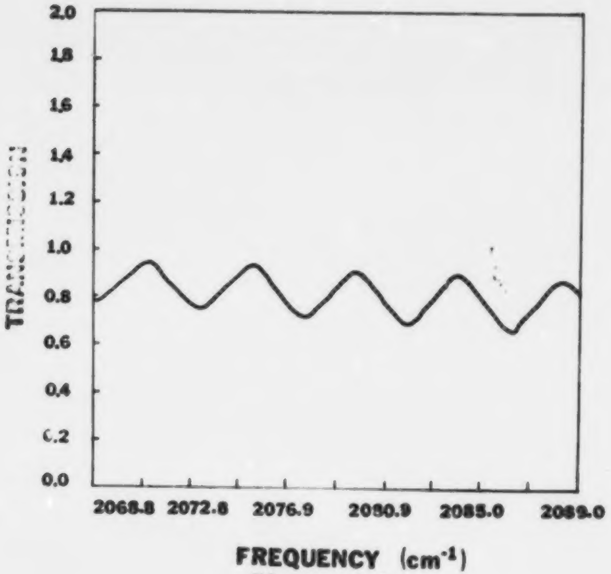


Figure 4.8

CALCULATED SPECTRUM, CO,
RESOLUTION 2.3 cm^{-1}
(Compare with F4.5)

The spectra shown in F4.9 (a-c) were reconstructed from EPA ROSE data of a combination of SO_2 (7.56×10^{18} molecules cm^{-2}) and CH_4 (1.23×10^{19} molecules cm^{-2}) with a spectral slit width of 1 cm^{-1}. Figure 4.10 (a-c) shows the computed spectra for the same regions with the same amounts of gas and the same slit width used for the calculations. The most obvious disagreement appears between the curves of F4.9b (experimental) and F4.10b. It would appear that some gas other than SO_2 and CH_4 (or noise) (see 4.4) was in the sample to produce the additional structure. An examination of the entire series shows a general agreement, but again the experimental curves show an indication of a higher resolution than that used for the computed spectra. It would also appear that the use of a lower value of concentration would produce a computed spectra which would come closer to matching the experimental curves. Probably both a lower concentration and a higher resolution would produce a spectrum which would agree with the experimental spectrum.

This discussion illustrates the necessity for an accurate normalization procedure and precision in determining instrumental resolution to avoid ambiguity in spectral analysis. The resolution of spectrometers generally changes over their spectral range. This makes it necessary that either the experimenter or the analyst must be able to compensate for the effect.

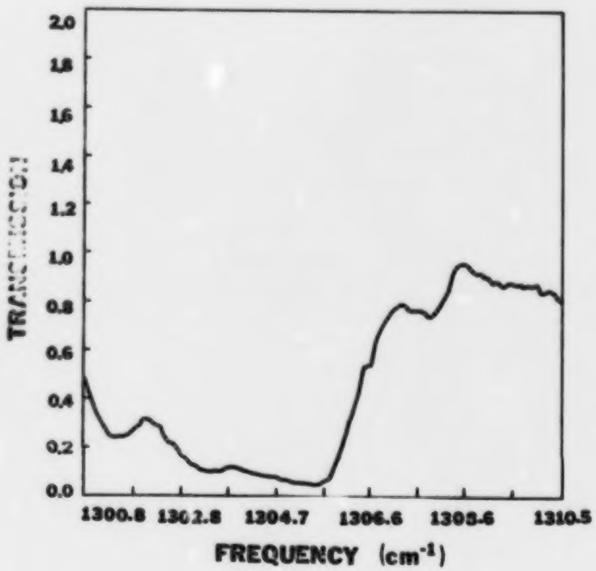


Figure 4.9a

EXPERIMENTAL SPECTRUM OF
COMBINATION OF SO_2 AND CH_4;
NOMINAL SLIT WIDTH 1.0 cm^{-1}
(ROSE SYSTEM) (1300.8-1310.5 cm^{-1})
(Compare with F4.10a,b,c)

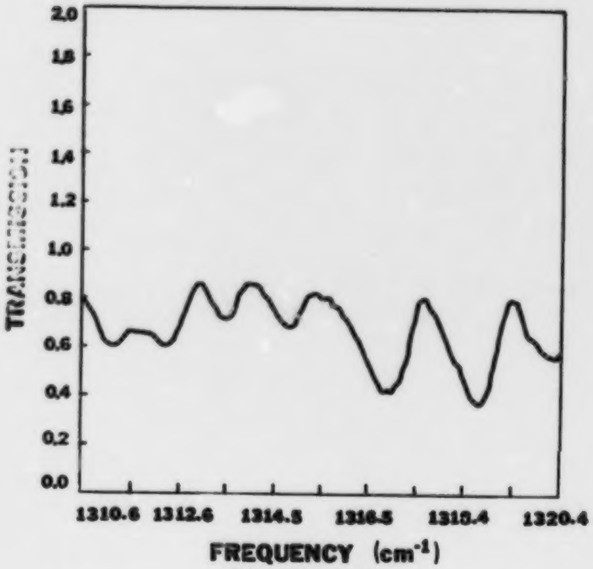


Figure 4.9b

EXPERIMENTAL SPECTRUM OF
COMBINATION OF SO_2 AND CH_4;
NOMINAL SLIT WIDTH 1.0 cm^{-1}
(ROSE SYSTEM) (1310.6-1320.4 cm^{-1})
(Compare with F4.10a,b,c)

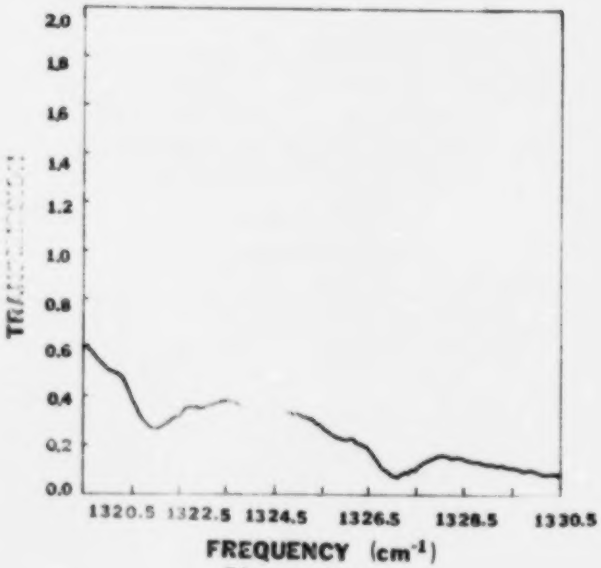


Figure 4.9c

EXPERIMENTAL SPECTRUM OF
COMBINATION OF SO_2 AND CH_4,
NOMINAL SLIT WIDTH 1.0 cm^{-1}
(ROSE SYSTEM) (1320.5-1330.5 cm^{-1})
(Compare with F4.10a,b,c)

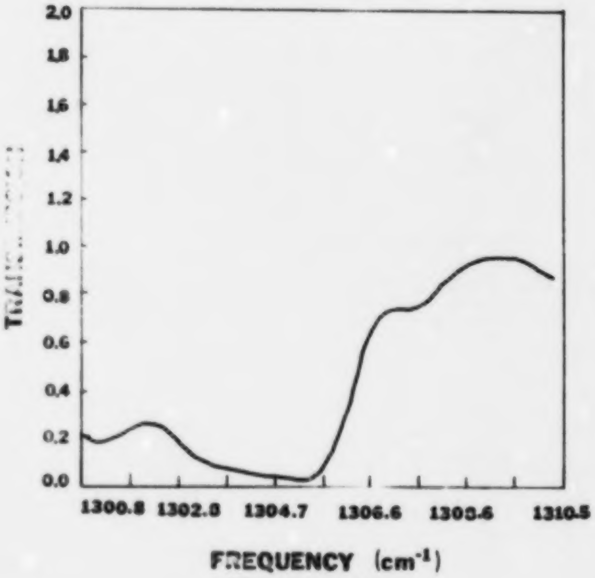


Figure 4.10a

CALCULATED SPECTRUM, SO_2 AND CH_4
(RESOLUTION 1.0 cm^{-1})
(1300.8-1310.5 cm^{-1})
(Compare with F4.9[illegible],b,c)

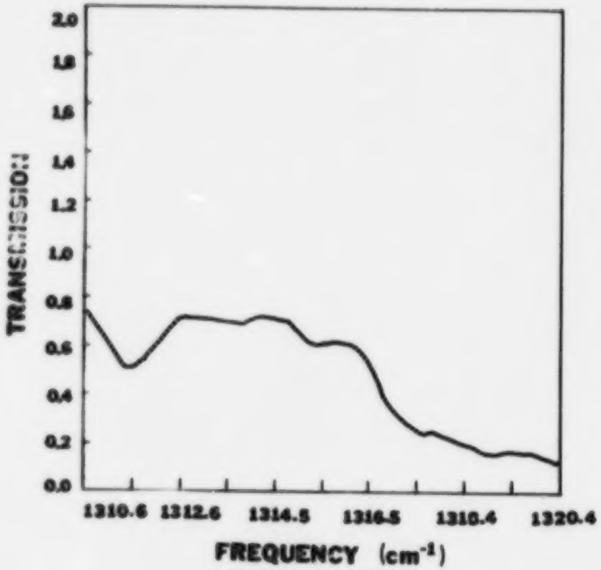


Figure 4.10b

CALCULATED SPECTRUM, SO_2 AND CH_4
(RESOLUTION 1.0 cm^{-1})
(1310.6-1320.4 cm^{-1})
(Compare with F4.9a,b,c)

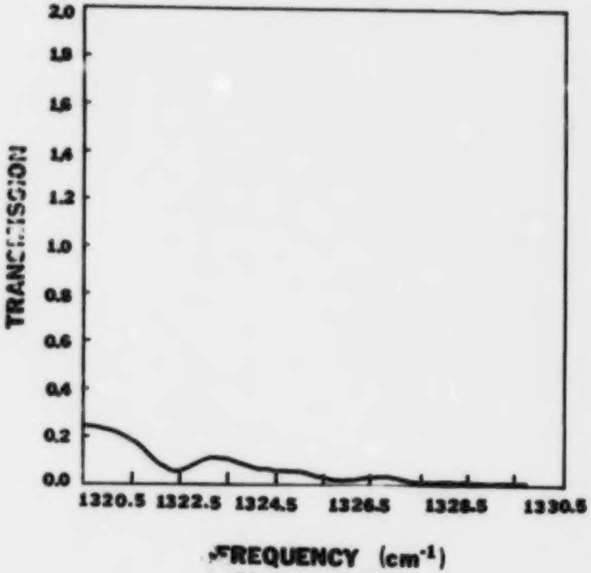


Figure 4.10c

CALCULATED SPECTRUM, SO_2 AND CH_4
(RESOLUTION 1.0 cm^{-1})
(1320.5-1330.5 cm^{-1})
(Compare with F4.9a,b,c)

4.2 CALCULATION OF THE CONCENTRATION OF CO

A composite spectrum of H_2O, CO_2, CO, CH_4, and SO_2 was prepared on the ROSE spectrophotometer over the spectral region from 1250 to 1330 cm^{-1} and in the range 2050 to 2150 cm^{-1}. Because only a relative intensity was obtained, normalization proceeded by obtaining a background run containing only spectra of H_2O and CO_2. The ratio of the two spectra gave the normalized spectrum I/I_o. The EPAGAS program was then used to calculate the concentration of CO in three spectra regions, 2073-2094, 2094-2114, and 2114-2132 cm^{-1}. T4.1 shows the results of the calculation. The spectral regions are listed in the second column. Several initial trial concentrations were used to determine the mathematical stability of the process. As can be seen for each region, different starting estimates made little change in final calculated concentrations. However, as discussed in 4.1, the resolution used in the spectral library strongly affects the accuracy. Part of the calculations summarized in T4.1 used the nominal spectral library resolution (3.0 cm^{-1}) for this region. However, as pointed out in 4.1, the resolution used in preparing the data was actually closer to 2.3 cm^{-1}. Nine values of concentration were calculated assuming a resolution of 2.3 cm^{-1}. These values average 1.26×10^{19} molecules cm^{-2}; the sample concentration measured chemically was 1.5×10^{19} molecules cm^{-2}. It is believed that this agreement is excellent in view of the calibration and normalization problems.

Note that the spectral resolution was carefully determined in the region 2073-2094 cm^{-1} and that the errors are the least there, increasing as the calculation is performed for more distant regions.

By using program DEGRADE (Deutschman and Calfee, 1967) before the fitting routine in EPAGAS, the actual resolution of the sample was determined to be 2.3 cm^{-1}. This choice of resolution was verified by the results given in T4.1.

Table 4.1

CALCULATION OF THE CONCENTRATION OF A SAMPLE OF CO BY EPAGAS TO AN EXPERIMENTAL SPECTRUM AT SEVERAL SPECTRAL REGIONS AND TWO INSTRUMENTAL RESOLUTIONS

Code	Region Spectral	Frequency Correction Δν	Resolution	CO Trial (Est.) Concentration (molecules cm^{-2})	CO Calculated Concentration (molecules cm^{-2})	Final Transmissions	Averages for a Given Resolution and Region Cal. Conc.	Error
1a	2073.0-2094.0	0.0	3.0	1.50x10^{19}	7.33x10^{18}	2.51		
b	2073.0-2094.0	0.0	3.0	1.88x10^{19}	7.32x10^{18}	2.51	7.33x10^{18}	2.51
c	2073.0-2094.0	0.0	2.3	5.38x10^{18}	1.18x10^{19}	1.25		
d	2073.0-2094.0	0.0	2.3	1.50x10^{19}	1.18x10^{19}	1.25		
e	2073.0-2094.0	0.0	2.3	1.88x10^{18}	1.18x10^{19}	1.25	1.18x10^{19}	1.25
2a	2094.0-2114.0	0.0	3.0	1.50x10^{19}	9.27x10^{18}	2.24		
b	2094.0-2114.0	0.0	3.0	1.88x10^{18}	9.29x10^{18}	2.24	9.28x10^{18}	2.24
c	2094.0-2114.0	0.0	2.3	5.38x10^{18}	1.26x10^{19}	2.18		
d	2094.0-2114.0	0.0	2.3	1.50x10^{19}	1.26x10^{19}	2.18		
e	2094.0-2114.0	0.0	2.3	1.88x10^{19}	1.26x10^{19}	2.18	1.26x10^{19}	2.18
3a	2114.0-2134.0	0.0	3.0	1.50x10^{19}	8.71x10^{18}	1.96		
b	2114.0-2134.0	0.0	3.0	1.88x10^{19}	8.71x10^{18}	1.96	8.71x10^{18}	1.96
c	2114.0-2134.0	0.0	2.3	5.38x10^{18}	1.35x10^{19}	2.19		
d	2114.0-2134.0	0.0	2.3	1.50x10^{19}	1.35x10^{19}	2.19		
e	2114.0-2.34.0	0.0	2.3	1.88x10^{19}	1.35x10^{19}	2.19	1.35x10^{19}	2.19

Average CO Concentration and Error for 3.0 resolution: 8.44x10^{18}; Error 2.24.

Average CO Concentration and Error for 2.3 resolution: 1.26x10^{19}; Error 1.87.

Sample Concentration: 1.50x10^{18} mol cm^{-2} (Obtained by Chemical Analysis).

The calculated concentration with the least error (1.26x10^{19}) compares well with the sample concentration (1.50x10^{18}).

NOTE: Powers of ten suppressed in errors quoted above. Concentrations are in molecules cm-2.

Figures 4.11 and 4.12 show, respectively, the data produced by the ROSE spectrophotometer and the best fitting calculated spectrum. The slight differences in this case are probably due to noise. Although the shapes of these spectra match quite well, the absolute transmission differences account for the slight discrepancy in calculated and experimental concentration. These absolute transmission differences are unresolvable without better normalization.

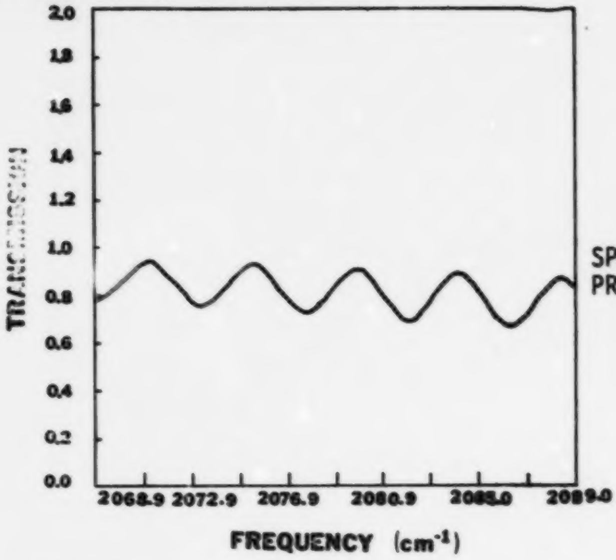


Figure 4.11

SPECTRUM OF CO CALCULATED BY PROGRAM DEGRADE (RESOLUTION 2.3 cm^{-1})
(Compare with F4.12)

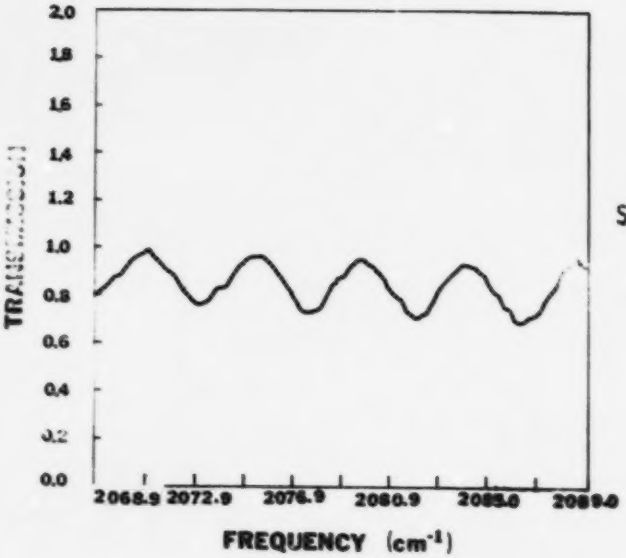


Figure 4.12

SPECTRUM OF CO, EXPERIMENTAL DATA FROM ROSE SPECTROPHOTOMETER
(Compare with F4.11)

The effect of resolution on the goodness of fit in EPAGAS may be seen in F4.13. In the first of these, the dashed curve represents the given experimental data, the full line is the calculated value. The results of the calculation are given in (resolution 3.0 cm^{-1}) T4.1. These curves should be contrasted with the first calculation listed in T4.1 using 2.3 cm^{-1} (line 7), shown in F4.14. The much better fit of given and calculated data results in a more accurate value of concentration and once again shows the importance of resolution matching.

The efficiency of the program in performing a "best fit" may be seen in F4.14. Notice that the two curves are not sufficiently alike to be fitted exactly. However, the optimization of the adjustable parameters of the calculated spectrum has clearly resulted in an excellent fit in a least-squares sense. An understanding of the change of a typical spectrum with a small change of concentration may be seen in F4.15. Clearly good accuracy and low noise of the instrument data are required for accurate analysis.

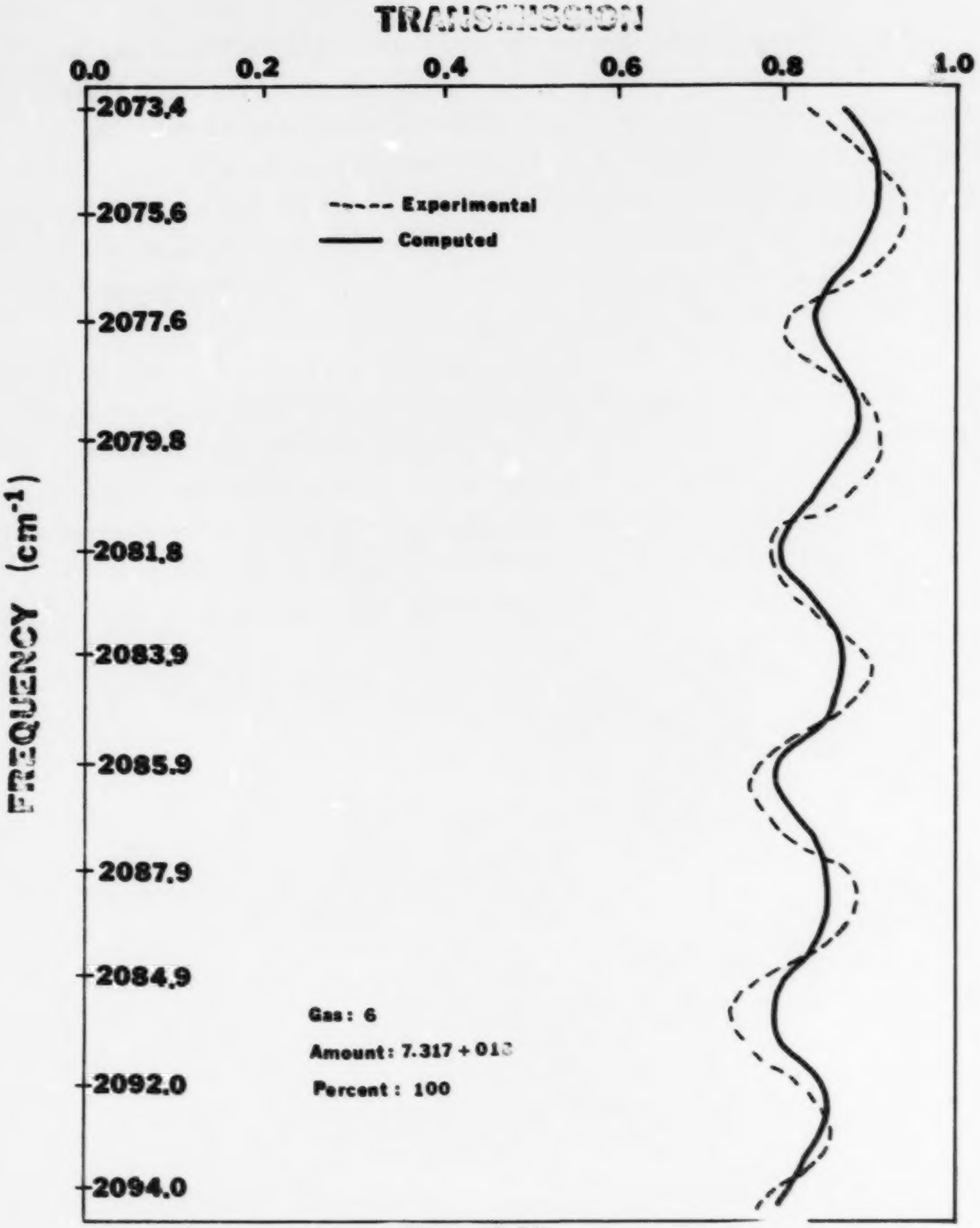


Figure 4.13

COMPUTED BEST-FIT TRANSMISSION OF CO (EPAGAS PROGRAM)
RESOLUTION 3.0 cm^{-1}

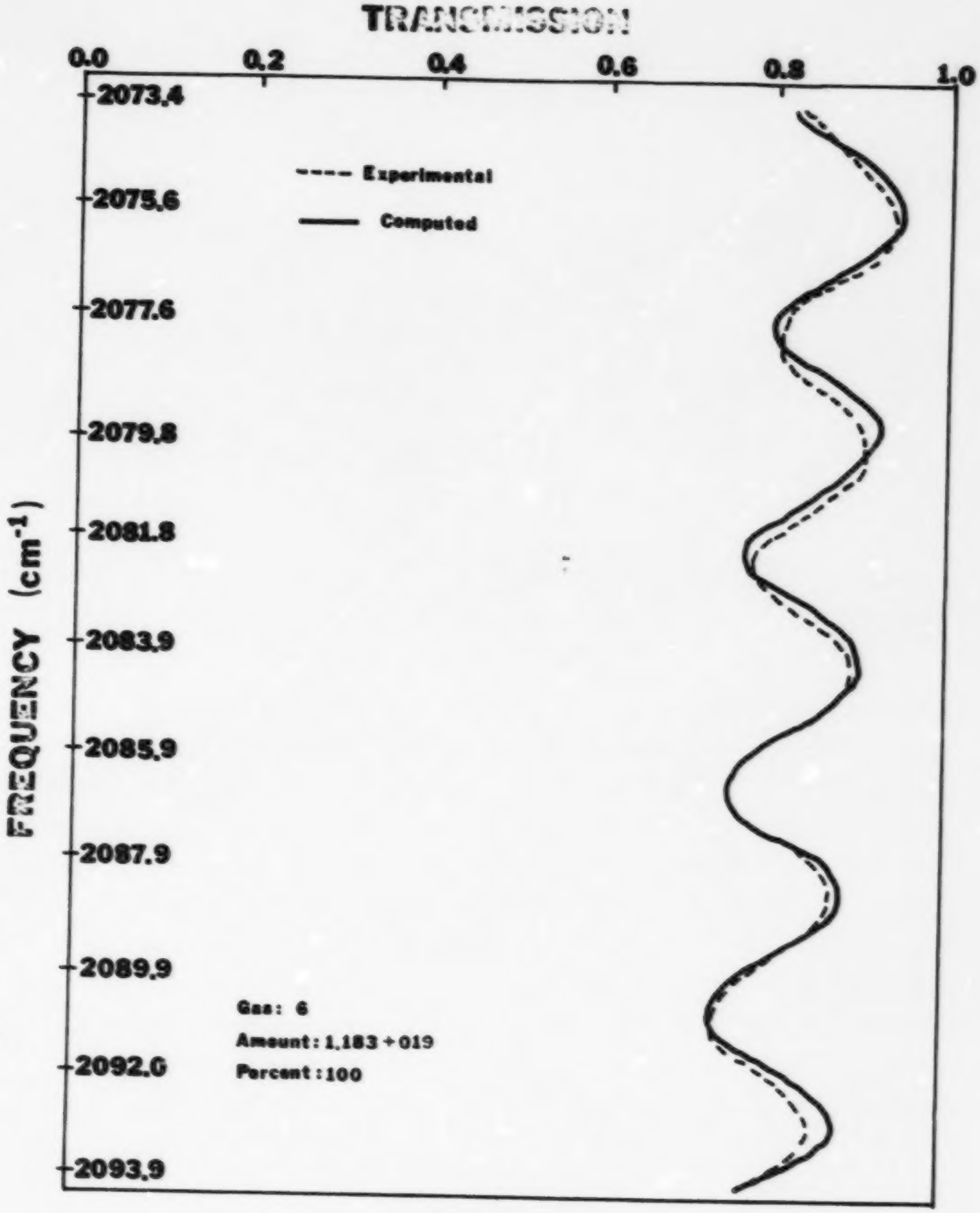


Figure 4.14

COMPUTED BEST-FIT TRANSMISSION OF CO (EPAGAS PROGRAM)
RESOLUTION 2.3 cm^{-1}

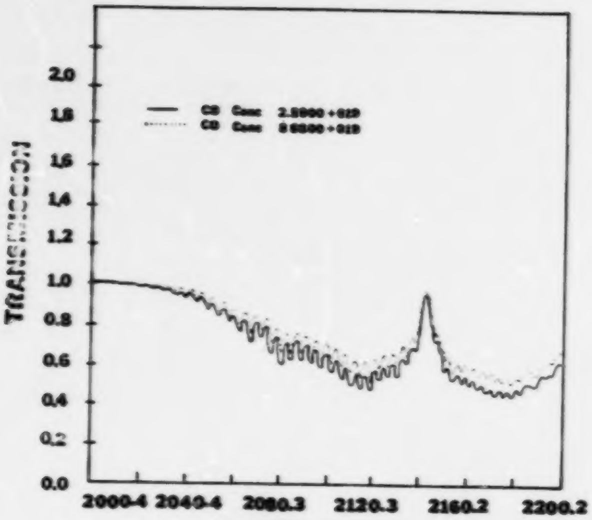


Figure 4.15

CHANGE IN SPECTRUM OF CO
FOR SMALL CHANGE OF CONCENTRATION
(Calculated by Program DEGRADE)

4.3 CALCULATION OF THE CONCENTRATION OF SO_2 IN THE PRESENCE OF CH_4

The gas sample cell used in obtaining data for CO_2, as discussed in 4.2, contained also SO_2 and CH_4. These gases have significant spectra in the region from 1150 to 1420 cm^{-1} (see F1.1 (a-c)). However, the data have good signal-to-noise ratio only in the ranges 1300-1330 cm^{-1} and 1250-1300 cm^{-1}. We first consider the 1300-1330 cm^{-1} range. In this range, the spectrophotometer data shows evidence of incomplete elimination of water spectra or noise, hence the data from the Fourier transform spectrometer were utilized. In this range, the spectral structure of CH_4 does not lend itself to fitting uniquely, and we here seek only to find the concentration of SO_2 in the presence of interference due to CH_4.

Table 4.2 shows the results of concentration calculations by MINMYZD for several different conditions of initial trial concentration and frequency corrections. The stability of the calculated concentration of SO_2, regardless of starting point, is apparent. The frequency corrections listed for the various regions are linear frequency shifts due to an uncertainty of the zero point in the Fourier transform.

As expected, the concentrations calculated for CH_4 here vary widely and are untrustworthy. However, the presence of CH_4 does not radically affect the value of SO_2 concentration computed in any of the spectral regions. The average value of SO_2 concentration molecules cm^{-2} is very close to the chemically determined concentration in spite of the many uncertainties in the data parameters. It is particularly encouraging to note that the concentration of the sample lay outside the span of library concentrations, but the program EPAGAS, by use of the subroutine FINDTRN, was able to extrapolate properly. The library does not include such large concentrations because it is not expected to be so large in field experiments, even in highly polluted geographical regions.

4.4 CALCULATION OF THE CONCENTRATION OF CH_4

An attempt was made to analyze the spectra for the concentration of CH_4. The results of this calculation were not as successful as the determination of the concentrations of CO and SO_2. The cause of the discrepancy is either noise or unknown interfering spectra (or both) in the data. The problem is illustrated in F4.16. There we see that the optimum fit of CH_4 to the data still allows large errors because structure is found in the data which does not come from the CH_4 molecule.

The source of the structure very likely lies among the following:

a. The library spectra of CH_4 is incorrect.

b. There may be small SO_2 lines from an adjacent band not found in the library.

c. The sample contained an unknown gas.

d. The spectrum contains instrumental noise.

e. Water lines are incompletely cancelled.

Table 4.2

CALCULATION OF THE CONCENTRATION OF THE SAMPLE OF SO_2 IN THE PRESENCE OF CH_4 BY APPLICATION OF EPAGAS TO AN EXPERIMENTAL SPECTRUM

Code	Spectral Region	Frequency Correction $\Delta\nu$	Trial (Est.) Concentrations (molecules cm^{-2})		Calculated Concentrations (molecules cm^{-2})		Final Transmission Error	Averages for a Given Frequency and Range		
			CH_4	SO_2	CH_4	SO_2		CH_4	SO_2	Error
1a*	1300.0-1310.0	0.0	$7.56x10^{18}$	$1.23x10^{19}$	$5.44x10^{18}$	$9.20x10^{19}$	24.89			
b	1300.0-1310.0	0.0	$1.23x10^{19}$	$7.56x10^{18}$	$6.33x10^{18}$	$8.37x10^{18}$	31.77	$6.33x10^{18}$	$8.37x10^{18}$	31.77
c*	1300.0-1310.0	-0.9	$7.56x10^{18}$	$1.23x10^{19}$	$7.98x10^{18}$	$1.47x10^{19}$	18.61			
d	1300.0-1310.0	-0.9	$1.23x10^{19}$	$7.56x10^{18}$	$7.98x10^{18}$	$9.47x10^{18}$	18.67	$7.98x10^{18}$	$9.47x10^{18}$	18.67
e	1300.0-1310.0	-1.2	$7.56x10^{18}$	$1.23x10^{19}$	$7.98x10^{18}$	$6.77x10^{18}$	14.61			
f	1300.0-1310.0	-1.2	$1.23x10^{19}$	$7.56x10^{18}$	$8.00x10^{18}$	$7.10x10^{18}$	14.61	$7.99x10^{18}$	$6.94x10^{18}$	14.61
2a*	1310.0-1320.0	0.0	$7.56x10^{18}$	$1.23x10^{19}$	$-6.54x10^{17}$	$7.77x10^{18}$	7.93			
b*	1310.0-1320.0	0.0	$1.23x10^{19}$	$7.56x10^{18}$	$-6.54x10^{17}$	$7.77x10^{18}$	7.93	-	-	-
c	1310.0-1320.0	-0.9	$7.56x10^{18}$	$1.23x10^{19}$	$6.68x10^{18}$	$6.86x10^{18}$	4.87			
d	1310.0-1320.0	-0.9	$1.23x10^{19}$	$7.56x10^{18}$	$6.75x10^{18}$	$6.84x10^{18}$	4.87	$6.72x10^{18}$	$6.85x10^{18}$	4.87
e	1310.0-1320.0	-1.2	$7.56x10^{18}$	$1.23x10^{19}$	$7.53x10^{18}$	$6.90x10^{18}$	4.89			
f	1310.0-1320.0	-1.2	$1.23x10^{19}$	$7.56x10^{18}$	$7.53x10^{18}$	$6.89x10^{18}$	4.89	$7.53x10^{18}$	$6.90x10^{18}$	4.89
3a	1320.0-1330.0	0.0	$7.56x10^{18}$	$1.23x10^{19}$	$3.15x10^{18}$	$6.09x10^{18}$	9.98			
b	1320.0-1330.0	0.0	$1.23x10^{19}$	$7.56x10^{18}$	$3.15x10^{18}$	$6.09x10^{18}$	9.98	$3.15x10^{18}$	$6.09x10^{18}$	9.98
c	1320.0-1330.0	-0.9	$7.56x10^{18}$	$1.23x10^{19}$	$1.89x10^{19}$	$6.25x10^{18}$	6.42			
d	1320.0-1330.0	-0.9	$1.23x10^{19}$	$7.56x10^{18}$	$2.27x10^{19}$	$6.23x10^{18}$	6.37	$2.08x10^{19}$	$6.24x10^{18}$	6.40
e	1320.0-1330.0	-1.2	$7.56x10^{18}$	$1.23x10^{19}$	$1.99x10^{19}$	$6.36x10^{18}$	5.69			
f	1320.0-1330.0	-1.2	$1.23x10^{19}$	$7.56x10^{18}$	$2.65x10^{19}$	$6.37x10^{18}$	5.51	$2.32x10^{19}$	$6.37x10^{18}$	5.60

Table 4.2 (cont.)

Code	Spectral Region	Frequency Correction $\Delta\nu$	Trial (Est.) Concentrations (molecules cm^{-2}) CH_4	Trial (Est.) Concentrations (molecules cm^{-2}) SO_2	Calculated Concentrations (molecules cm^{-2}) CH_4	Calculated Concentrations (molecules cm^{-2}) SO_2	Final Transmission Error	Averages for a Given Frequency and Range CH_4	Averages for a Given Frequency and Range SO_2	Error
4a	1300.0-1330.0	0.0	$7.56x10^{18}$	$1.23x10^{19}$	$5.11x10^{18}$	$6.54x10^{18}$	20.05			
b	1300.0-1330.0	0.0	$1.23x10^{19}$	$7.56x10^{18}$	$5.18x10^{18}$	$6.56x10^{18}$	20.05	$5.15x10^{18}$	$6.55x10^{18}$	20.05
c	1300.0-1330.0	-0.9	$7.56x10^{18}$	$1.23x10^{19}$	$8.00x10^{18}$	$6.84x10^{18}$	11.49			
d	1300.0-1330.0	-0.9	$1.23x10^{19}$	$7.56x10^{18}$	$8.00x10^{18}$	$6.84x10^{18}$	11.49	$8.00x10^{18}$	$6.84x10^{18}$	11.49
e	1300.0-1330.0	-1.2	$7.56x10^{18}$	$1.23x10^{19}$	$8.05x10^{18}$	$7.04x10^{18}$	9.22			
f	1300.0-1330.0	-1.2	$1.23x10^{19}$	$7.56x10^{18}$	$8.05x10^{18}$	$7.04x10^{18}$	9.22	$8.05x10^{18}$	$7.04x10^{18}$	9.22
5a*	1310.0-1330.0	0.0	$7.56x10^{18}$	$1.23x10^{19}$	$1.71x10^{18}$	$6.20x10^{18}$	9.77			
b*	1310.0-1330.0	0.0	$1.23x10^{19}$	$7.56x10^{18}$	$1.91x10^{18}$	$6.20x10^{18}$	9.77	-	-	-
c	1310.0-1330.0	-0.9	$7.56x10^{18}$	$1.23x10^{19}$	$9.62x10^{18}$	$6.38x10^{18}$	6.34			
d	1310.0-1330.0	-0.9	$1.23x10^{19}$	$7.56x10^{18}$	$9.62x10^{18}$	$6.38x10^{18}$	6.34	$9.62x10^{18}$	$6.38x10^{18}$	6.34
e	1310.0-1330.0	-1.2	$7.56x10^{18}$	$1.23x10^{19}$	$1.06x10^{19}$	$6.52x10^{18}$	6.05			
f	1310.0-1330.0	-1.2	$1.23x10^{19}$	$7.56x10^{18}$	$1.08x10^{19}$	$6.52x10^{18}$	6.05	$1.07x10^{19}$	$6.52x10^{18}$	6.05

Average Concentrations for all Regions	Resolution	CH_4	SO_2	Error
	0.0	$4.88x10^{18}$	$7.00x10^{18}$	20.60
	-0.9	$6.88x10^{18}$	$7.16x10^{18}$	9.55
	-1.2	$5.39x10^{18}$	$6.75x10^{18}$	8.07

Sample Concentrations: $CH_4 = 1.23x10^{19}$ molecules cm^{-2} (Obtained by Chemical Analysis)
$SO_2 = 7.56x10^{18}$ molecules cm^{-2}

* Not used in Averages

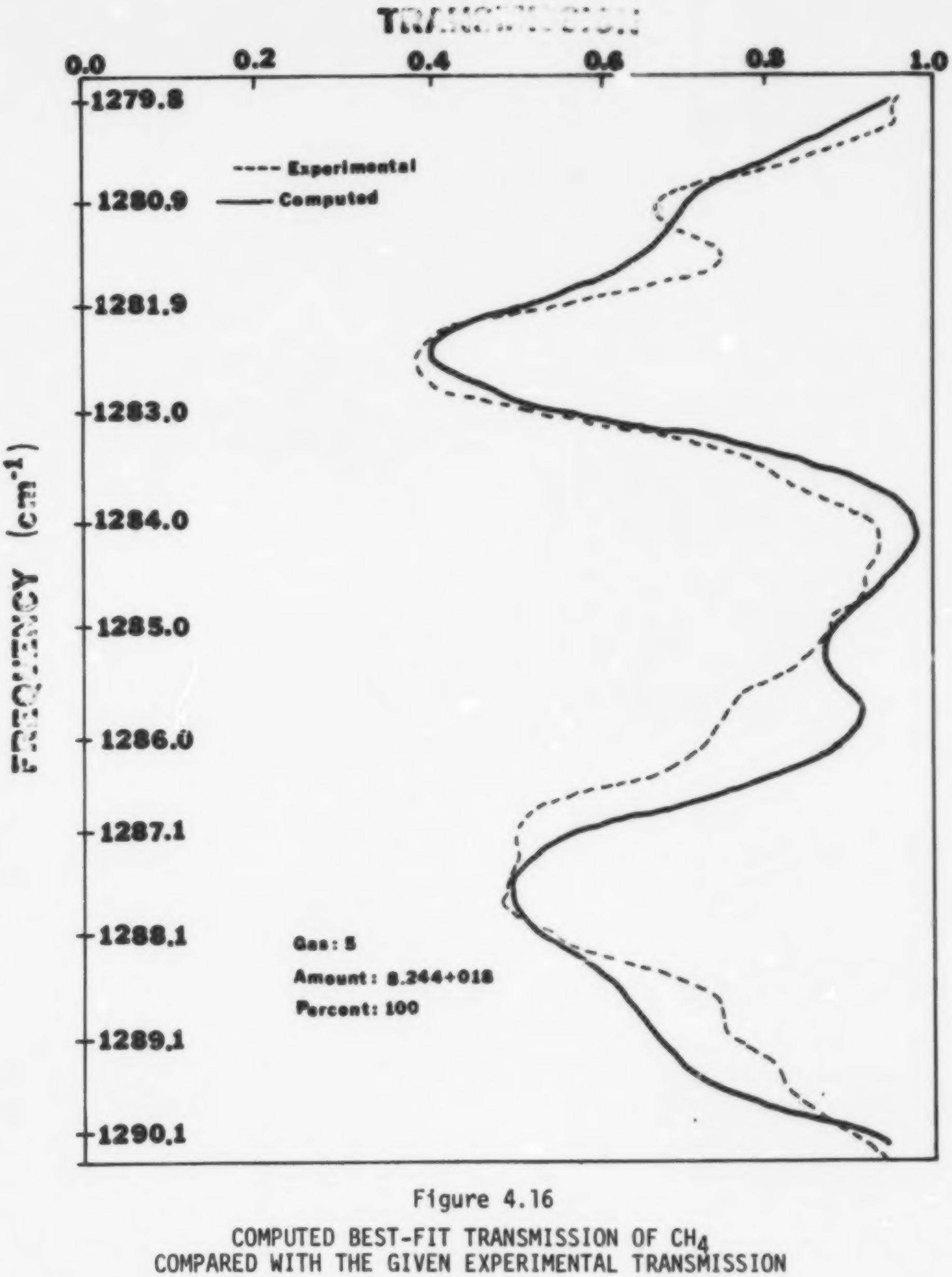


Figure 4.16

COMPUTED BEST-FIT TRANSMISSION OF CH_4
COMPARED WITH THE GIVEN EXPERIMENTAL TRANSMISSION

We comment on these possibilities:

a. The library spectrum of CH_4 has been checked by many experimenters. It is possible, but considered unlikely, that it is incorrect.

b. A chemical analysis by EPA of the sample did not indicate any other gases. The analysis method should be reviewed to see if it would reveal other gases than those known to be in the sample.

c. The spectrum furnished by EPA from the ROSE spectrophotometer clearly has noise in some regions of the spectrum. The section analyzed here appeared less noisy than any other by visual inspection, but if the noise were low frequency drifts, it would not be seen as obvious noise.

d. An initial suspicion that water absorptions were responsible was discounted by the investigation described below.

In summary, the reason for the discrepancy between the value of 1.23×10^{19} molecules per cm^2 of CH_4 and the value of 8.22×10^{18} calculated by EPAGAS is not known at this time. Both instrumental uncertainties and possible library inadequacies must be investigated. It should be noted that even with these difficulties, the values differ by a factor of only 1.5.

Table 4.3 shows a summary of calculations done to determine the concentrations of CH_4 and H_2O in the spectrum. The water values are low (for water) and are scattered, when compared in different frequency regions, indicating that the spectra contains very little water absorption. The CH_4 values are, on the other hand, quite consistent. Figure 4.16 shows the impossibility of fitting the given spectrum with a CH_4-only absorption spectrum. Structure in the given spectrum cannot be matched by any concentration of the CH_4 spectrum.

Table 4.3

CALCULATION OF THE CONCENTRATION OF A SAMPLE OF CH_4 AND H_2O BY APPLICATION OF EPAGAS TO AN EXPERIMENTAL SPECTRUM

Code	Region Spectral	Frequency Correction $\Delta\nu$	Trial (Est.) Concentrations (molecules cm^{-2}) H_2O	Trial (Est.) Concentrations (molecules cm^{-2}) CH_4	Calculated Concentrations H_2O	Calculated Concentrations CH_4	Final Transmission Error
1a	1250.0-1260.0	-0.6	3.73×10^{22}	1.23×10^{19}	-2.17×10^{21}	7.79×10^{18}	6.07
b	1260.0-1270.0	-0.6	3.73×10^{22}	1.23×10^{19}	6.82×10^{20}	8.46×10^{18}	6.88
c	1270.0-1280.0	-0.6	3.73×10^{22}	1.23×10^{19}	-1.44×10^{20}	8.67×10^{18}	6.73
d	1280.0-1290.0	-0.6	3.73×10^{22}	1.23×10^{19}	5.04×10^{20}	7.96×10^{18}	7.14
2a	1250.0-1260.0	-0.9	3.73×10^{22}	1.23×10^{19}	-9.14×10^{20}	8.05×10^{18}	5.41
b	1260.0-1270.0	-0.9	3.73×10^{22}	1.23×10^{19}	7.09×10^{20}	8.17×10^{18}	6.14
c	1270.0-1280.0	-0.9	3.73×10^{22}	1.23×10^{19}	4.42×10^{20}	6.85×10^{18}	8.68
d	1280.0-1290.0	-0.9	3.73×10^{22}	1.23×10^{19}	1.80×10^{20}	7.90×10^{18}	10.20
		$\Delta\nu$			H_2O	CH_4	Error
Averages for all regions and a given frequency Correction		-0.6			-2.82×10^{20}	8.22×10^{18}	6.71
		-0.9			1.04×10^{20}	7.74×10^{18}	7.61

5. CONCLUSIONS AND RECOMMENDATIONS

Many geophysical studies require the measurement of constituent concentrations in situations where the identifying signals from many components are mixed in the spectrometer. In remote measurement of aerosols by electromagnetic scatter, in measurement of ocean constituents, in the measurement of pollutants in the water and in the air, the same complex combination of identifying spectra occur. Means of separating and measuring a large number of components of a mixture usually require analysis by a large computer because the data set is quite large. Not only must the field data be large enough to give sufficient statistically independent information, but the identifiers (spectra) require a large computer memory to avoid breaking the problem into too many pieces.

The calculational algorithms utilized in the program EPAGAS have been shown in this report to determine efficiently and accurately the concentrations of gas from a complex long-path infrared spectrum. Two important requirements are:

1. The spectral library must be accurate in wavelength, pressure and temperature broadening parameters, and spectral line or band intensity.
2. The spectrophotometer used to obtain the field data must accurately measure the spectral intensity (at least relatively), and the wavelength and the spectral resolution must match the resolution of the spectra in the library.

The first requirement is met (with qualifications detailed in A2) by the spectral library assembled for this report (F1.1(a-c)), but adequate spectral data analysis of complex spectra is rare in the scientific literature, as discussed in A1. Other gases than those considered in this report are important for pollution monitoring. The technological changes in the United States resulting from the fuel shortage and the expected mineral shortage will require new industrial developments such as oil shale processing. These can be expected to result in new effluents whose spectra will most likely not be well known. Indeed, only a small portion of atmospheric pollutants presently in industrial areas are well enough known spectroscopically.

Thus, it is emphatically recommended that continuing support be given to competent spectroscopists to obtain the basic spectroscopic data needed for air and water pollution. Absorption spectra, Raman spectra, and emission spectra are needed for the various techniques presently in development.

It is further recommended that vigorous efforts be made to obtain field data under controlled conditions to develop experience with the whole system. Initially, the experiments must include, along the path, point samplers to establish the levels of gaseous pollutants for comparison with the spectral analysis. Remote sensing systems must not be considered complete until proven in comparison with more conventional methods.

6. ACKNOWLEDGMENTS

The authors would like to thank Ms. K. Yoshida and Mr. D. Gregory for their meticulous assistance in processing the data for this report. Mr. R. Lawrence has provided wise counsel in the use of the MINMYZD program. Dr. R. Schwiesow has made essential contributions to the library of spectra and intelligent aid in all phases of the task. Dr. W. Herget of EPA (Research Triangle Park) has clarified many important problems and supplied experimental data. The authors express their gratitude to Ms. B. Stahl for her skill and monumental patience in producing the manuscript.

7. LIST OF SYMBOLS

I	Radiation Intensity
K'	Absorption coefficient (cm^{-1})
K	Absorption coefficient (cm^2 molecules^{-1})
W	Area density (or concentration) (cm^{-2})
P	Density absorbing gas per unit area of radiation (cm^{-3})
N	Noise power
ν	Spatial frequency (wavenumber) (cm^{-1})
T	Temperature (^{o}K)
P	Pressure (dynes-cm^{-2} or atm)
$S(\nu)$	Spectrophotometer output
E	Error function
Wg	Given concentration (molecules cm^{-2})
Wc	Computed concentration (molecules cm^{-2})
CF	Confidence factor
λ	Wavelength (cm)
C	Speed of light in vacuum (cm sec^{-1})
f	Frequency (sec^{-1})
S/N	Signal-to-noise ratio

8. REFERENCES

Abels, L. L., and J. H. Shaw (1966), Widths and strengths of vibration - rotation lines in the fundamental band of nitric oxide, J. Mol. Spectry., 20, 11-28.

Box, G. E. P., and Mervin E. Muller (1958), A note on the generation of random normal deviates, Ann. Math. Statist., 29, 610-611.

Burch, D. E. (1970), Investigation of the absorption of infrared radiation by atmospheric gases, Semiannual Technical Report, Jul. 31, 1969 - Jan. 31, 1970, Philco Ford Corp., Aeronutronic Div., Newport Beach, CA., Contract F19628-69-C-0263, U4784.

Burriss, R. C., E. Griggs, T. Connolly, and W. J. Veigele (1972), Land Use Planning for Air Quality in the Pikes Peak Area, Kaman Sciences Corp., Colorado Springs, CO, Contract USE-CO-08-00-0002, K-72-946-FR.

Calfee, R. F. (1966), Infrared absorption properties of atmospheric gases, J. Quant. Spectrosc. Radiat. Transfer, 6, 221-228.

Calfee, R. F. (1971), A note on terminologies used in gaseous absorption processes, NOAA Tech. Rept. ERL 211-WPL15, Boulder, CO.

Cossee, P., and J. H. Schachtschneider (1966), Vibrational analysis of acetone, acetaldehyde, and formaldehyde, J. Chem. Phys., 44, (1), 97-111.

[illegible], E. M., and R. F. Calfee (1967), Two computer programs to produce theoretical absorption spectra of water vapor and carbon dioxide, ESSA Tech. Rept., IER 31-ITSA 31, Boulder, CO.

Lawrence, R. S., and M. Hallenbeck (1965), A method for obtaining the parameters of electron-density profiles from topside ionograms, NBS Tech. Note No. 315, Boulder, CO.

Levin, A., and C. F. Meyer, The infrared absorption spectra of acetylene, ethylene and ethane, J. Opt. Soc. Am., 16, (3), 137-164.

Matthes, W. (1973), "Unfolding of Composite Spectra by Linear Regression," Private communication of a paper to be published.

McClatchey, R. A., R. W. Fenn, F. E. Volz, J. S. Garing, and J. E. A. Selby (1971), Optical properties of the atmosphere, *AFCRL Environmental Research Paper #354*.

McClatchey, R. A., W. S. Benedict, S. A. Clough, D. E. Burch, R. F. Calfee, and K. Fox (1973), AFCRL atmospheric absorption line parameters compilation, *AFCRL Environmental Research Paper #434*, AFCRL-TR-73-0096, A.F. Sys. Comm., Hanscom Field, Bedford, MA.

Milne, J. B., and A. Ruoff (1967), Coriolis coupling constants in sulfur trioxide, *J. Mol. Spectry.*, *23*, 408-415.

Nakagawa T., and V. Morino (1971), Coriolis Interactions in the ν_4 and ν_6 band of formaldehyde, *J. Mol. Spectros.*, *38*, 84-106.

Nielsen, H. H., and E. F. Barker (1931), Infrared absorption bands in hydrogen sulphide, *Phys. Rev.* *37*, 727-732.

Pierson, R. H., A. N. Fletcher, and E. St. C. Gantz (1956), Catalog of infrared spectra for qualitative analysis of gases, *Anal. Chem.*, *28*, 1218-1239.

Shapiro, M. M., and H. P. Gush (1966), The collision-induced fundamental and first overtone bands of oxygen and nitrogen, *Can. J. Phys.*, *44*, 949-963.

Shaw, J. H. (1956), Nitric oxide fundamentals, *J. Chem. Phys.*, *24*, (2), 399-402.

Slutz, R. J., and J. R. Winkelman (1964), Shape of the magnetospheric boundary under solar wind pressure, *J. Geophys. Res.*, *69*, (23), 4933-4948.

Smith, L. G. (1940), Fine structure of the 3.35 μ band of ethylene, *J. Chem. Phys.*, *8*, 798-799.

Smith, W. L., and I. M. Mills (1964), Coriolis perturbation in the infrared spectrum of ethylene, *J. Chem. Phys.*, *40*, (8), 2095-2109.

Streigg, M. L., and C. R. Claysmith (1972), Design and construction of a system for remote optical sensing of emissions, *Final Rept.*, Contract EPA-CPA 22-69-142, General Dynamics, Convair Aerospace Div., San Diego, CA.

U. S. Naval Res. Lab., *American Petroleum Institute Research Project 44*, (in-house project).

Westwater, E. R., and A. Cohen (1973), Application of Backus-Gilbert inversion technique to determination of aerosol size distributions from optical scattering measurement, *Appl. Opt.*, *12*, (6), 1340-1348.

APPENDIX 1

TABLE OF CONTENTS

Page

A1. Sources and Contents of the Library of Spectra 71
  A1.1 Motivation for the Library 71
  A1.2 Sources and Content 71
  A1.3 Programs Used in Processing Spectra 91

APPENDIX 1

A1. SOURCES AND CONTENTS OF THE LIBRARY OF SPECTRA

A1.1 MOTIVATION FOR THE LIBRARY

As outlined in the body of the report and explained in detail in A2, EPAGAS is a program which employs a least-squares fitting method to match an experimentally taken spectrum, say of a polluted atmosphere, with an internally generated reference spectrum. Through constructing reference spectra by adjusting the concentrations of gases in its library, EPAGAS fits the data spectrum to a reference spectrum and thus determines the concentration of the gases in the data spectrum.

The closer that the library comes to containing spectra of all the molecules found in the experimental spectra, the more accurate are the results of EPAGAS. (See 3.1.3.) This is the motivation for assembling a large and accurate library of spectra.

Physically, the library consists of a set of condensed computer cards which contain values of the absolute transmittance of each gas at each concentration at each frequency used by EPAGAS-a total of about 256,000 points.

A1.2 SOURCES AND CONTENT

Spectra for the 10 gases shown in F1.1 were assembed by this laboratory* over the past 10 years and are summarized in TA1.1. The spectra of the first seven gases listed in the table (set numbers 1A1 through 7B1) were calculated theoretically using line parameters and self-broadening constants. SO_2, HNO_3, and NH_3 spectra were taken experimentally. All assumed room temperature (296°K).

---

*Wave Propagation Laboratory, ERL/NOAA. These data were assembled from theoretical computations and experimental data from many sources. See McClatchey et al., 1973.

Table A1.1

SPECTRA FROM THE WPL DATA SET

SET NUMBER	GAS	WAVENUMBER REGION	INCREMENT (cm^{-1})	RESOLUTION	NUMBER OF POINTS PER CONCENTRATION	CONCENTRATIONS (molecules cm^{-2})
1A1	H_2O-Water	740.0-1450.0	0.2	1.0	3551	7.46×10^{21}, 2.24×10^{22}, 3.73×10^{22}, 5.22×10^{22}, 6.77×10^{22}
1B1	H_2O-Water	1800.0-3200.4	0.6	3.0	2335	7.46×10^{20}, 2.24×10^{21}, 3.73×10^{21}, 5.22×10^{21}, 6.77×10^{21}
2A1	CO_2-Carbon Dioxide	710.0-1100.0	0.2	1.0	1951	8.61×10^{20}, 1.08×10^{21}, 1.29×10^{21}
2B1	CO_2-Carbon Dioxide	1810.2-2499.0	0.6	3.0	1149	8.61×10^{20}, 1.08×10^{21}, 1.29×10^{21}
3A1	O_3-Ozone	970.0-1100.0	0.2	1.0	651	1.08×10^{17}, 1.88×10^{17}, 2.69×10^{17}, 5.38×10^{17}, 1.08×10^{18}
3B1	O_3-Ozone	2040.0-2139.0	0.6	3.0	166	1.08×10^{17}, 1.88×10^{17}, 2.69×10^{17}, 5.38×10^{17} 1.08×10^{18}
4A1	N_2O-Nitrous Oxide	1120.0-1360.0	0.2	1.0	1201	5.38×10^{17}, 1.08×10^{18}, 1.88×10^{18}
4B1	N_2O-Nitrous Oxide	2130.6-2599.8	0.6	3.0	783	5.38×10^{17}, 1.08×10^{18}, 1.88×10^{18}
5A1	CH_4-Methane	1180.0-1400.0	0.2	1.0	1101	2.69×10^{18}, 5.38×10^{18}, 1.08×10^{19}, 1.88×10^{19}, 2.69×10^{19}
5B1	CH_4-Methane	2488.2-3199.8	0.6	3.0	1187	2.69×10^{18}, 5.38×10^{18}, 1.08×10^{19}, 1.88×10^{19}, 2.69×10^{19}
6B1	CO-Carbon Monoxide	2000.4-2259.6	0.6	3.0	433	5.38×10^{18}, 1.08×10^{19}, 1.88×10^{19}, 2.69×10^{19}, 5.38×10^{19}
7B1	HCL-Hydrochloric Acid	2400.6-3169.8	0.6	3.0	1283	2.69×10^{16}, 1.08×10^{17}, 2.69×10^{17}, 1.08×10^{18}, 2.69×10^{18}
20.1	SO_2-Sulfur Dioxide	1074.8-1231.2	0.2	1.0	783	1.06×10^{18}, 3.54×10^{18}, 1.06×10^{19}, 3.54×10^{19}
20.2	SO_2-Sulfur Dioxide	1299.2-1409.2	0.2	1.0	551	3.54×10^{16}, 1.06×10^{17}, 3.54×10^{17}, 1.06×10^{18}, 3.54×10^{18}
30.1	HNO_3-Nitric Acid	824.8- 951.2	0.2	1.0	633	3.54×10^{17}, 1.06×10^{18}, 3.54×10^{18}
30.2	HNO_3-Nitric Acid	1234.8-1383.8	0.2	1.0	746	1.06×10^{17}, 3.54×10^{17}, 1.06×10^{18}, 3.54×10^{18}
40.1	NH_3-Ammonia	1112.0-1211.4	0.2	1.0	498	3.54×10^{17}, 1.06×10^{18}, 3.54×10^{18}, 1.06×10^{19}

The theoretical gas spectra were calculated using program DEGRADE (Deutschman and Calfee, 1967) for specific resolutions, temperatures, concentrations, and data increments. In other applications, it is possible to calculate new theoretical spectra for differing sets of these parameters. For the remaining experimental spectra in TA1.1 and all of those in TA1.2, except NO, it is only possible to generate lower resolution spectra of different concentrations using programs LOWRES and MORTRAN. See A1.4 for brief explanations and listings of these and other programs used in processing the library of spectra.

In TA1.1, "increment" refers to the spacing on the wavenumber axis between transmittance values. Wavenumber ν is related to frequency f and wavelength λ by $\nu=\frac{f}{c}=\frac{1}{\lambda}$, where c is the speed of light. Note that in the lower region, 700 to 1450 cm^{-1}, the increments are 0.2 cm^{-1}. In the upper region, 1800 to 3200 cm^{-1}, the increments are 0.6 cm^{-1} and they are spaced evenly from 1800.0 cm^{-1}. This standardization of spacing is necessitated by the manner in which EPAGAS compares spectra. In the lower region, most spectra in the library were of 1.0 cm^{-1} resolution, while in the upper region, they were typically of 3.0 cm^{-1} resolution. Thus the 0.2 and 0.6 cm^{-1} spacings give five data points per resolution element, providing essentially complete information about the spectra.*

---

*Trying to compare higher or lower resolution experimental spectra with library spectra will introduce errors, but effects on EPAGAS outputs are presently unknown. It should be noted that most commercial spectrometers change resolution while making a large frequency scan. The adjustments to the program can be made quite simply by changing the input data to correspond to the library requirements.

To increase the library, the literature was searched for usable spectra of 23 molecules. These molecules have been found in polluted atmospheres, but no high resolution spectra had been collected for use in this application. Usable data for only 11 gases, those listed in TA1.2, could be found with the time and resources available. For the remaining 12 gases, either no spectra were available or published information was unusable because no absolute scale of transmittance or absorption was given or because the concentrations could neither be calculated nor estimated with available information. Except in a few special cases (SO_2), it was not possible to obtain spectra experimentally due to the expense and time involved with the procurement and handling of samples and the proper ratioing of data to obtain absolute values of transmittance.

A summary sheet for each of the 11 molecules for which useful information was found follows this discussion of general techniques (A1.3). Note that gas 4 is missing. This gas (SO_3) was deleted, after the coding system had been finalized, when it was determined that the spectra were in error. The important parameters required for the conversion of the published data to a form usable by program EPAGAS are listed in seven columns. First, a photocopy of the region of interest (column 1) was made from the literature and photo-enlarged to a size suitable for scaling on a digitizer. Absorption values for even increments of wavenumber (column 4) were punched onto cards, which in turn were used to produce one card for each frequency together with its corresponding value of transmittance, plus peripheral identifiers such as the resolution (column 2) and concentration (column 3) found in the literature.

It was necessary to produce these cards to manipulate the spectra with existing in-house programs. SPTR, a simple read-punch program, was used for this purpose. As needed, the spectra were then interpolated linearly to finer spacings (column 5) using program INTERP and/or degraded to a lower resolution (column 6) using program LOWRES. The latter program employs a sliding triangular slit function as a low pass filter to smooth high-resolution data into low-resolution data.

Table A1.2

SPECTRA PROCESSED FROM THE SCIENTIFIC LITERATURE

SET NUMBER	GAS	WAVENUMBER REGION (cm^{-1})	INCREMENT (cm^{-1})	NUMBER OF POINTS PER CONCENTRATION	LOWEST CONCENTRATION (molecules cm^{-2})
1	NO-Nitric Oxide	1805.4 - 1935.0	0.6	217	1.06×10^{18}
2	H_2S-Hydrogen Sulfide	2623.8 - 2755.2	0.6	220	1.06×10^{20}
3a	C_2H_4-Ethylene	800.0 - 1098.0	0.2	1491	1.06×10^{18}
3b	C_2H_4-Ethylene	1380.0 - 1450.0	0.2	351	3.54×10^{18}
3c	C_2H_4-Ethylene	1800.0 - 1955.4	0.6	260	3.54×10^{18}
3d	C_2H_4-Ethylene	1970.4 - 2101.8	0.6	220	3.54×10^{19}
3e	C_2H_4-Ethylene	2936.4 - 3049.8	0.6	190	3.54×10^{18}
3f	C_2H_4-Ethylene	3050.4 - 3200.4	0.6	251	1.06×10^{18}
5	NO_2-Nitrogen Dioxide	2853.0 - 2967.0	0.6	191	3.54×10^{17}
6a	CH_3CH_2HCO-Acetone	700.0 - 1149.8	0.2	2250	1.06×10^{18}
6b	CH_3CH_2HCO-Acetone	1150.0 - 1315.8	0.2	830	3.54×10^{17}
6c	CH_3CH_2HCO-Acetone	1325.0 - 1450.0	0.2	626	3.54×10^{17}
6d	CH_3CH_2HCO-Acetone	2900.4 - 3099.0	0.6	332	1.06×10^{18}
7b	H_2CO-Formaldehyde	851.0 - 1071.6	0.2	1104	1.06×10^{19}
7c	H_2CO-Formaldehyde	1077.0 - 1237.0	0.2	801	3.54×10^{18}
7e	H_2CO-Formaldehyde	1239.0 - 1439.0	0.2	1001	3.54×10^{18}
8a	C_2H_6-Ethane	792.6 - 870.0	0.2	388	3.54×10^{18}
8b	C_2H_6-Ethane	1360.0 - 1450.0	0.2	451	3.54×10^{18}
8c	C_2H_6-Ethane	2863.2 - 3050.4	0.6	313	3.54×10^{17}

Table A1.2 (cont.)

SET NUMBER	GAS	WAVENUMBER REGION (cm^{-1})	INCREMENT (cm^{-1})	NUMBER OF POINTS PER CONCENTRATION	TEST CONCENTRATION (molecules cm^{-2})
9a	n -C_4H_{10}-Normal Butane	700.0 - 1450.0	0.2	3751	3.54×10^{18}
9b	n -C_4H_{10}-Normal Butane	1800.0 - 3200.4	0.6	2335	3.54×10^{18}
10a	i -C_4H_{10}-Iso-Butane	700.0 - 1450.0	0.2	3751	1.06×10^{19}
10b	i -C_4H_{10}-Iso-Butane	1800.0 - 3200.4	0.6	2335	1.06×10^{19}
11a	C_3H_8-Propane	700.0 - 1450.0	0.2	3751	1.06×10^{19}
11b	C_3H_8-Propane	1800.0 - 3200.4	0.6	2335	1.06×10^{19}
12a	C_5H_{12}-Pentane	700.0 - 1450.0	0.2	3751	3.54×10^{18}
12b	C_5H_{12}-Pentane	1800.0 - 1999.8	0.6	334	1.06×10^{19}
12c	C_5H_{12}-Pentane	2544.0 - 3200.4	0.6	1095	3.54×10^{18}

Note: Insufficient data was found for O_2, C_6H_6, SO_3, HCN, H_2O_2, C_3H_4O, C_2H_4O, C_4H_8O, C_3H_8, C_7H_8, and N_2-N_2 collisional spectra. In most cases, the spectrum was of too low resolution or unknown concentration.

Finally program MORTRAN, using Bouguer-Beer's law, calculated transmittance values for five concentrations from each data point for storage in the library. Any particular problems encountered or special techniques employed for a molecule are summarized in the final paragraph(s) on the individual summary sheets.

In several cases, molecular bands were found that could not be processed, either because absolute transmittance scales were not used by the author, resolution was extremely low, or no value of concentration could be deduced. Such bands were listed as missing if other bands of the same molecule could be processed. It is possible, though unlikely, that bands exist which the literature search did not find. Such bands were neither processed nor listed as missing.

The accuracy of the spectra in the library was controlled primarily by the accuracy of the blowup process. Typically, a 2 x 6 in. journal spectrum was enlarged to 18 x 54 in. and data points were taken every 0.04 in. on the frequency axis of the enlargement. The uncertainty in reproducing the position of a spectrum on the frequency axis was about 0.12 in., or ±3 of the data-taking intervals (column 4). The uncertainty in controlling the length of the transmittance axis is ±5 percent of full scale. Both of these uncertainties assume the reported spectra were ratioed and calibrated.

The spectra for several gases at concentrations expected in the atmosphere were nearly straight lines. That is, more than 90 percent of the band had transmittance values above 0.97. Such spectra are nearly useless for use in EPAGAS. Thus, the first concentration which caused at least 10 percent of the band to fall under 97 percent transmission, and the next four higher concentrations in the standard sequence were calculated and used in the library. The standard sequence is . . . , 1.06×10^{n}, 3.54×10^{n}, $1.06 \times 10^{n+1}$, $3.54 \times 10^{n+1}$, $1.06 \times 10^{n+2}$, . . . (molecules cm^{-2}), where n is an integer. The "lowest concentration" in the last column on table A1.2 refers to the first number selected from this sequence and implies the next four which were used in the library.

Concentrations on TA1.1 and TA1.2 and on the summary sheets are given in molecules per square centimeter for a 1-km path. 1.06×10^{18} molecules cm^{-2} corresponds to 3 torr partial pressure at 296°K. For conversion to other units, consult A4 (Calfee, 1971) which lists many of the standard units used in studying gaseous absorption processes and summarizes their interrelationships. See the reference for detailed explanations.

Figures 1.1b and 1.1c display selected spectra for each of the gases processed. The overlays to F1.1a and F1.1b depict band designations. Where disagreement occurs between literature sources, the band was not identified, but the disagreement was noted on the summary sheet. Due to the complicated vibration and rotation modes of the larger molecules in F1.1c, individual bands are not easily identifiable and therefore no overlay is provided.

Figure A1.1 summarizes the state of knowledge in the infrared of absorbing gases found in the atmosphere.

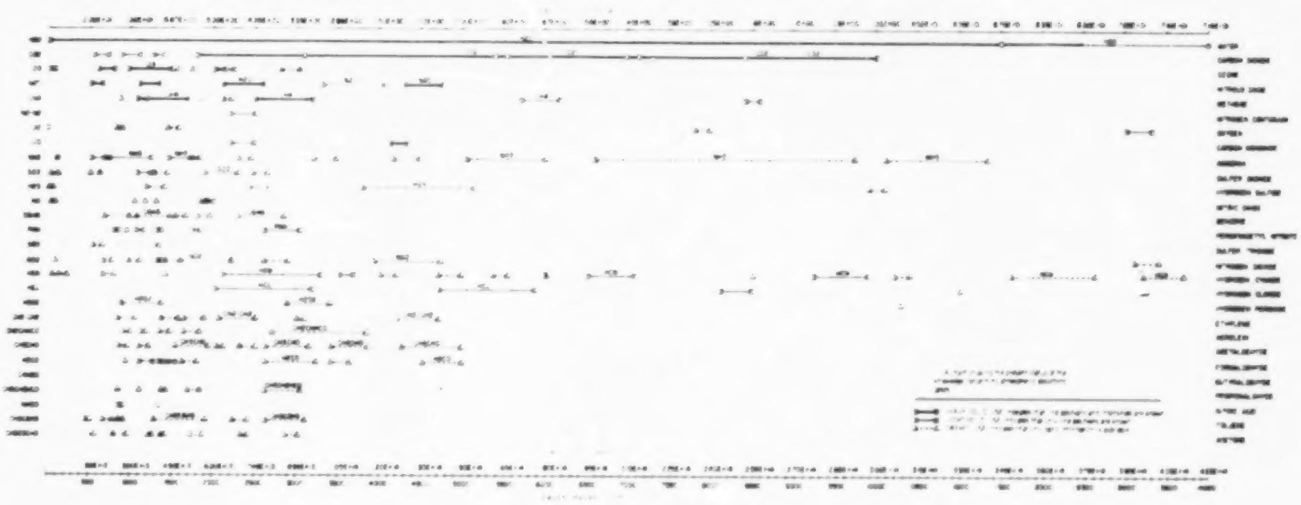


Figure A1.1

INFRARED ABSORBING REGIONS OF ATMOSPHERIC GASES

GAS 1: NO-Nitric Oxide

1 Frequency Range (cm^{-1})	2 Original Resolution (cm^{-1})	3 Original Concentration (mol. cm^{-2})	4 Increment Data Taken (cm^{-1})	5 Converted Increment (cm^{-1})	6 Converted Resolution (cm^{-1})	7 Interpolated Concentrations (mol. cm^{-2})	Source
1805.4-1935.0	-	-	-	0.6	3.0	1.06×10^{18}-1.06×10^{20}	1,2

1. J. Chem. Phys., Vol. 24, #2, p. 399, 1956, Shaw.
2. J. Molec. Spec., Vol. 20, p. 11, 1966, Abels & Shaw.

The NO spectra were calculated from theoretical line positions (Shaw, 1956), theoretical line strengths (Abels & Shaw, 1966), and experimental half-widths (Abels & Shaw, 1966) using program DEGRADE. Where not reported, positions were computed using formulas presented in Shaw, 1956, and half widths were estimated from adjacent lines. Line parameter cards containing this unreported information are identifiable by a missing last significant figure in the field describing the position and/or half-width.

GAS 2: H_2S-Hydrogen Sulfide

1 Frequency Range (cm^{-1})	2 Original Resolution (cm^{-1})	3 Original Concentration (mol. cm^{-2})	4 Increment Data Taken (cm^{-1})	5 Converted Increment (cm^{-1})	6 Converted Resolution (cm^{-1})	7 Interpolated Concentrations (mol. cm^{-2})	Source
2623.8-2755.2	1.2	2.7×10^{20}	0.6	-	-	1.06×10^{20}-1.06×10^{22}	1

1. Phys. Review, Vol. 37, p. 728, 1931, Nielsen & Barker

A strong band, 1100-1400cm^{-1}, is missing. The resolution claimed by Nielsen and Barker seems high. The apparent resolution of their spectrum is close to 3.0cm^{-1}, the resolution required for this region. Therefore, LOWRES was not used in processing this spectrum for the library.

GAS 3: C_2H_4-Ethylene

	1 Frequency Range (cm^{-1})	2 Original Resolution (cm^{-1})	3 Original Concentration (mol. cm^{-2})	4 Increment Data Taken (cm^{-1})	5 Converted Increment (cm^{-1})	6 Converted Resolution (cm^{-1})	7 Interpolated Concentrations (mol. cm^{-2})	Source
a	800.0-1098.0	0.8	1.1×10^{19}	0.4	0.2	-	1.06×10^{18}-1.06×10^{20}	1
b	1380.0-1450.0	1.5	3.0×10^{19}	0.2	-	-	3.54×10^{18}-3.54×10^{20}	2
c	1800.0-1955.4	1.6	1.0×10^{20}	0.6	-	-	3.54×10^{18}-3.54×10^{20}	2
d	1970.4-2101.8	1.5	2.5×10^{20}	0.6	-	-	3.54×10^{19}-3.54×10^{21}	2
e	2936.4-3049.8	0.6	3.4×10^{19}	0.6	-	3.0	3.54×10^{18}-3.54×10^{20}	3
f	3050.4-3200.4	1.5	2.4×10^{19}	0.6	-	-	1.06×10^{18}-1.06×10^{20}	2

1. J. Chem. Phys., Vol. 40, #8, p. 2096, 1964 Smith & Mills.
2. J. Opt. Soc. Amer., Vol. 16, #3, p. 155, 1928, Levin & Meyer.
3. J. Chem. Phys., Vol. 8, p. 799, 1940, Smith.

No accurate concentrations were reported by Levin and Meyer. The original concentrations given in column 3 were first estimated from Sadier low resolution band strengths and then refined by comparison with neighboring spectra where possible.

For instance, the original spectra in the 2936-3049 cm^{-1} and 3050-3200cm^{-1} regions were assumed to have been taken at concentrations of 2.7×10^{19} mol-cm^2 and 3.0×10^{19} mol-cm^2, respectively. When converted to a concentration of 3.54×10^{18} mol-cm^{-2}, a noticeable mismatch occurred between the two spectra at 3050cm^{-1}. To cause the two spectra to coincide at 3050cm^{-1}, the assumed concentration for the first spectrum had to be increased to 3.4×10^{19} mol-cm^{-2} (25.3% increase) and the second lowered to 2.4×10^{19} mol-cm^{-2} (19.1% decrease). These refined figures were the ones reported on the summary sheet, and they reflect a 50% uncertainty in the reported concentration.

Once again, the reported resolution seems too high in all the Levin and Meyer spectra. Apparent resolution is about 3.0cm^{-1}. Note that the apparent resolution is satisfactory for the 1800-3200cm^{-1} region but too low for the 700-1500cm^{-1} region.

GAS 5: NO_2-Nitrogen Dioxide

1 Frequency Range (cm^{-1})	2 Original Resolution (cm^{-1})	3 Original Concentration (mol. cm^{-2})	4 Increment Data Taken (cm^{-1})	5 Converted Increment (cm^{-1})	6 Converted Resolution (cm^{-1})	7 Interpolated Concentrations (mol. cm^{-2})	Source
2853.0-2967.0	0.6	1.1×10^{19}	0.6	-	1.0	3.54×10^{17}-3.54×10^{19}	1

---

1. Private communication from Goldman, University of Denver, 1969.

Only about one half of the band is available from the band center to one wing. Another band, centered at 1617cm^{-1} is available, but was omitted from the library due to strong H_2O absorption in this region.

GAS 6: CH_3CH_2HCO-Acetone

	1 Frequency Range (cm^{-1})	2 Original Resolution (cm^{-1})	3 Original Concentration (mol. cm^{-2})	4 Increment Data Taken (cm^{-1})	5 Converted Increment (cm^{-1})	6 Converted Resolution (cm^{-1})	7 Interpolated Concentrations (mol. cm^{-2})	Source
a	700.0-1149.8	(IR^{-7})	8.2×10^{19}	1.0	0.2	-	1.06×10^{18}-1.06×10^{20}	1
b	1150.0-1315.8	(IR^{-7})	3.2×10^{18}	1.0	0.2	-	3.54×10^{17}-3.54×10^{19}	1
c	1325.0-1450.0	(IR^{-7})	9.9×10^{18}	1.0	0.2	-	3.54×10^{17}-3.54×10^{19}	1
d	2900.4-3099.0	(IR^{-7})	9.9×10^{18}	1.0	0.6	-	1.06×10^{18}-1.06×10^{20}	1

---

1. J. Chem. Phys., Vol. 44, #1, p. 97, 1966, Cossee & Schachtschneider.

A band exists from 1680-1825cm^{-1} as reported by Cassee and Schachtschneider, but it was omitted due to strong H_2O absorption in this region.

The resolution is lower than desired (approximately 7cm^{-1}) due to the low resolution of the Beckman IR-7 spectrometer. Since a range of resolutions may be preset on this instrument, one needs to consult the workers to determine the exact resolution of the various runs.

GAS 7: H_2CO-Formaldehyde

	1 Frequency Range (cm^{-1})	2 Original Resolution (cm^{-1})	3 Original Concentration (mol. cm^{-2})	4 Increment Data Taken (cm^{-1})	5 Converted Increment (cm^{-1})	6 Converted Resolution (cm^{-1})	7 Interpolated Concentrations (mol. cm^{-2})	Source
a	851.0- 997.2	0.3	7.1×10^{19}	0.2	-	1.0	$1.06\text{x}10^{19}$-$1.06\text{x}10^{21}$	1
b	997.4-1071.6	0.3	3.5×10^{19}	0.2	-	1.0	$1.06\text{x}10^{19}$-$1.06\text{x}10^{21}$	1
c	1077.0-1237.0	0.3	4.7×10^{19}	0.2	-	1.0	$3.54\text{x}10^{18}$-$3.54\text{x}10^{20}$	1
d	1239.0-1325.0	0.3	2.4×10^{19}	0.2	-	1.0	$3.54\text{x}10^{18}$-$3.54\text{x}10^{20}$	1
e	1321.0-1439.0	0.3	5.3×10^{19}	0.2	-	1.0	$3.54\text{x}10^{18}$-$3.54\text{x}10^{20}$	1

---

1. J. Molec. Spec., Vol. 38, p. 90, 1971, Nakagawa & Morino.

Temperatures of the original spectra range from 150°C to 25°C, accounting for small discrepancies which may occur between regions. See Nakagawa & Morino for details.

Note an error on page 90, Figure le. The reported cell length is 100cm but should read 10cm.

Nakagawa and Morino report the band head at $1167cm^{-1}$ to be ν_4, while Herzberg claims it is ν_6.

GAS 8: C_2H_6-Ethane

	1 Frequency Range (cm^{-1})	2 Original Resolution (cm^{-1})	3 Original Concentration (mol. cm^{-2})	4 Increment Data Taken (cm^{-1})	5 Converted Increment (cm^{-1})	6 Converted Resolution (cm^{-1})	7 Interpolated Concentrations (mol. cm^{-2})	Source
a	792.6- 870.0	0.6	6.0×10^{19}	0.2	-	-	3.54×10^{18}-3.54×10^{20}	1
b	1360.0-1450.0	1.3	4.0×10^{19}	0.2	-	-	3.54×10^{18}-3.54×10^{20}	1
c	2863.2-3050.4	1.4	6.0×10^{19}	0.6	-	-	3.54×10^{17}-3.54×10^{19}	1

---

1. J. Opt. Soc. Amer., Vol. 16, #3, p. 155, 1928, Levin & Meyer.

No accurate concentrations were reported. The original concentrations given in column 2 were estimated from Sadler low resolution band strengths.

The reported resolution seems too high. The apparent resolution is about $1.0cm^{-1}$ in the 793-$870cm^{-1}$ region and about $3.0cm^{-1}$ in the remaining two regions.

GAS 9: C_4H_{10}-N-Butane

	1 Frequency Range (cm^{-1})	2 Original Resolution (cm^{-1})	3 Original Concentration (mol. cm^{-2})	4 Increment Data Taken (cm^{-1})	5 Converted Increment (cm^{-1})	6 Converted Resolution (cm^{-1})	7 Interpolated Concentrations (mol. cm^{-2})	Source
a	700.0-1450.0	-	2.5×10^{20}	5.0	0.2	-	3.54×10^{18}-3.54×10^{20}	1
b	1800.0-3200.4	-	2.5×10^{20}	5.0, 10.0	0.6	-	3.54×10^{18}-3.54×10^{20}	1

---

1. Anal. Chem., Vol. 28, p. 1218, 1956, Pierson, Fletcher, Grantz.

The resolution for all regions is unknown, but is approximately 10-15cm^{-1}. The instrument used was a Perkin-Elmer Model 21.

GAS 10: C_4H_{10}-I-Butane

	1 Frequency Range (cm^{-1})	2 Original Resolution (cm^{-1})	3 Original Concentration (mol. cm^{-2})	4 Increment Data Taken (cm^{-1})	5 Converted Increment (cm^{-1})	6 Converted Resolution (cm^{-1})	7 Interpolated Concentrations (mol. cm^{-2})	Source
a	700.0-1450.0	-	2.5×10^{20}	5.0	0.2	-	1.06×10^{19}-1.06×10^{21}	1
b	1800.0-3200.4	-	2.5×10^{20}	5.0, 10.0	0.6	-	1.06×10^{19}-1.06×10^{21}	1

---

1. Anal. Chem., Vol. 28, p. 1218, 1956, Pierson, Fletcher, Grantz.

The resolution for all regions is unknown, but is approximately 10-15cm^{-1}. The instrument used was a Perkin-Elmer Model 21.

GAS 11: C_3H_8-Propane

	1 Frequency Range (cm^{-1})	2 Original Resolution (cm^{-1})	3 Original Concentration (mol. cm^{-2})	4 Increment Data Taken (cm^{-1})	5 Converted Increment (cm^{-1})	6 Converted Resolution (cm^{-1})	7 Interpolated Concentrations (mol. cm^{-2})	Source
a	700.0-1450.0	-	2.5×10^{20}	5.0	0.2	-	1.06×10^{19}-1.06×10^{21}	1
b	1800.0-3200.4	-	2.5×10^{20}	5.0, 10.0	0.6	-	1.06×10^{19}-1.06×10^{21}	1

1. Anal. Chem., Vol. 28, p. 1218, 1956, Pierson, Fletcher, Grantz.

The resolution for all regions is unknown, but is approximately 10-15cm^{-1}. The instrument used was a Perkin-Elmer Model 21.

GAS 12: C_5H_{12}-Pentane

	1 Frequency Range (cm^{-1})	2 Original Resolution (cm^{-1})	3 Original Concentration (mol. cm^{-2})	4 Increment Data Taken (cm^{-1})	5 Converted Increment (cm^{-1})	6 Converted Resolution (cm^{-1})	7 Interpolated Concentrations (mol. cm^{-2})	Source
a	700.0-1450.0	-	9.1×10^{19}	5.0	0.2	-	3.54×10^{18}-3.54×10^{20}	1
b	1800.0-1999.8	-	9.1×10^{19}	5.0	0.2	-	1.06×10^{19}-1.06×10^{21}	1
c	2544.0-3200.4	-	9.1×10^{19}	10.0	0.6	-	3.54×10^{18}-3.54×10^{20}	1

---

1. U.S. Naval Research Lab., American Petroleum Institute Research Project 44.

The original spectrum for this molecule was for a liquid sample 0.174mm thick, and therefore it may differ significantly from the gaseous spectrum. Equivalent gaseous concentrations were figured from the density of liquid pentane and the thickness of the sample.

The resolution for all regions is unknown, but is approximately 10-15cm^{-1}.

A1.3 PROGRAMS USED IN PROCESSING SPECTRA

FORTRAN PROGRAM
CDC 3800

SPTR
M. J. Post
Aug. 17, 1973

Given absorption values A expressed as an integer number between 999 and 0, corresponding to 99.9 and 0.0 percent absorption, SPTR punches one card for each absorption value. On this card are punched wavenumber and the value of transmittance.

Input

Card A_i - A_n

Column	Field	Description
1-80	20I4	Absorption values A(I)

Output

Card A_i

Column	Field	Description
1-10	F10.2	Wavenumber V_i
11-20	F10.3	Transmittance T_i
54-58	F5.2	Resolution RES
59-70	E12.3	Concentration CONC
73-80	A8	Molecule name MOL

Note:
1. An EOF card is required at the end of the data cards.
2. RES, CONC, and MOL are entered on source cards, not data cards.
3. V_i are generated internally by changing source cards.

PROGRAM SPTR

THIS PROGRAM TAKES COMPACTED SPECTRAL ABSORPTION VALUES (FORMAT 20I4) AND OUTPUTS CARDS. ONE FREQUENCY AND ONE TRANSMISSION VALUE PER CARD, PLUS PERIPHERAL IDENTIFIERS.

IT IS USEFUL IN CONVERTING DIGITIZED ANALOG SPECTRAL INFORMATION (FROM STRIP CHARTS, FOR EXAMPLE) TO CARDS IN A FORMAT USED IN ASSOCIATED PROGRAMS.

```
      PROGRAM SPTR
      DIMENSION A(20)
    1 FORMAT (20F4.4)
    2 FORMAT (F10.2,F10.3,33X,F5.2,E12.3,2X,A8)
      V=1149.0
      DV=1.0
      MOL=8H ACETONE
      RES=0.0
      CONC=3.2E+18
    5 READ 1,(A(I),I=1,20)
      IF (EOF,60)20,10
   10 DO 15, K=1,20
      V=V+DV
      T=1.-A(K)
      PRINT 2,V,T,RES,CONC,MOL
   15 PUNCH 2,V,T,RES,CONC,MOL
      GO TO 5
   20 STOP
      END
```

FORTRAN PROGRAM
CDC 3800

LOWRES
M. J. Post
Aug. 17, 1973

Given a transmission spectrum T(N) of high resolution, LOWRES degrades it to a spectrum TF of lower resolution A.

Used in the computations are the given increment DV and the output increment DELV. A triangular slit function of half-width A is used to compute each new TF.

Input

Card A

Column	Field	Description
1-10	F10.3	Beginning wavenumber of given spectrum, V1
11-20	F10.3	End wavenumber of given spectrum, V2
21-25	F5.2	Increment given spectrum, DV
26-30	F5.2	Lower resolution desired, A
31-35	F5.2	Desired output increment, DELV

Card B

Column	Field	Description
11-20	F10.3	Given spectrum T(N)

Output

Card A_i

Column	Field	Description
1-10	F10.3	Wavenumber of output transmittance, VI_i
11-20	F10.5	Transmittance TF_i of output spectrum
54-58	F5.2	Resolution of output spectrum RES (=A)
50-70	E12.3	Concentration of output spectrum CONC
73-80	A8	Molecular name MOL

NOTE: An EOF card is required at the end of the data cards. DELV must be an exact multiple of DV and A must be an odd multiple of DV. CONC and MOL are entered on source cards, not data cards.

```
      PROGRAM LOWRES

      THIS PROGRAM TAKES A SPECTRUM OF HIGHER RESOLUTION GIVEN IN INCREMENTS
      OF DV AND DEGRADES IT TO A SPECTRUM OF RESOLUTION A, GIVEN IN INCREMENTS
      OF DELV, POSITIONED AT THE ORIGINAL FREQUENCIES.

      DELV MUST BE AN EXACT MULTIPLE OF DV.

      A MUST BE AN ODD MULTIPLE OF DV.

      DIMENSION T(2000)
    1 FORMAT (2F10.3,3F5.0)
    2 FORMAT (10X,F10.3)
    3 FORMAT (F10.3,F10.5,33X,F5.2,E12.3,2X,A8)
      RES=3.0
      CONC=1.05E+19
      MOL=8H      NO2
      READ 1,V1,V2,DV,A,DELV
      MULT=DELV/DV + .00001
      SLIT=A*2.
      NMX=SLIT/DV-.99999
      SLTFTR=DV/A**2
      VEND=V2-A+DV
      N=1
   20 READ 2,T(N)
      IF (EOF,60) 30,25
   25 N=N+1
      GO TO 20
   30 NJ=1
      VI=V1+A-DV
      V=VI+DV
   50 SUMT=0.
      DO 150 NN=NJ,NMX
      TT=(A-ABS(V-VI))* T(NN)
      SUMT=SUMT+TT
  150 V=V+DV
      TF=SUMT*SLTFTR
      PRINT 3,VI,TF,RES,CONC,MOL
      PUNCH 3,VI,TF,RES,CONC,MOL
      NJ=NJ+MULT
      NMX=NMX+MULT
      VI=VI+DV*MULT
      IF (VI.GT.VEND)  CALL EXIT
      V=VI-A+DV
      GO TO 50
      END
```

FORTRAN PROGRAM
CDC 3800

INTERP
M. J. Post
Aug. 17, 1973

Given a spectrum spaced in increments of DELV, INTERP linearly interpolates it to finer increments DIV and outputs it on cards.

Input

	Column	Field	Description
Card A			
	1-5	F5.2	Input spectral increment DELV
	6-10	F5.2	Output spectral increment DIV
Card B			
	1-10	F10.3	Beginning wavenumber of spectrum V1
	11-20	F10.5	First transmittance value of spectrum T1
	54-58	F5.3	Resolution of spectrum RES
	59-70	E12.3	Concentration of spectrum CONC
	73-80	A8	Molecular name MOL
Card C			
	11-20	F10.5	Transmittance values $T2_i$

Output

Card A_i

Column	Field	Description
1-10	F10.3	Wavenumber of transmittance value, VI_i
11-20	F10.5	Transmittance value T_i
54-58	F5.3	Resolution of spectrum RES
59-70	E12.3	Concentration of spectrum CONC
73-80	A8	Molecular name MOL

```
      PROGRAM INTERP
C
C     THIS PROGRAM TAKES A SPECTRUM SPACED IN INCREMENTS OF DELV AND LINEARLY
C     INTERPOLATES IT TO FINER INCREMENTS DIV.
C
C     DELV MUST BE AN INTEGRAL MULTIPLE OF DIV.
C
    1 FORMAT (2F5.2)
    2 FORMAT (10X,F10.5)
    3 FORMAT (F10.3,F10.5,33X,F5.3,F12.3,2X,A8)
      READ 1,DELV,DIV
      READ 3,V1,T1,RES,CONC,MOL
      N=DELV/DIV + .001
      VI=V1-DIV
   10 READ 2,T2
      IF (EOF,60) 20,12
   12 PART =(T2-T1)/N
      DO 15 K=1,N
      T=T1+PART*(K-1)
      VI=VI+DIV
      PUNCH 3,VI,T,RES,CONC,MOL
   15 PRINT 3,VI,T,RES,CONC,MOL
      T1=T2
      GO TO 10
   20 END
```

For a detailed discussion of program DEGRADE, see Deutschman and Calfee, 1967.

```
      PROGRAM DEGRADE
      DIMENSION W(5),F(5),AB(5),TT(5),TF(5),SUMT(5),TRAPSUM(5),FMT1(9),
     1 FMT2(9),FMT3(9),FMT4(9)
      COMMON/1/GNU(3000),SA(3000),AA (3000),CAY(1)/2/T(3000,5),I5
    1 FORMAT(9A8)
    2 FORMAT(F9.2,F12.6,10F10.5)
    3 FORMAT(/11H0PRESSURE =F9.5,3X8HSLIT A =F9.5,3X13HTEMPERATURE =F9.2
     1//12H ABSORPTANCE,9X13HCONCENTRATION/)
    4 FORMAT(24H KICKED OUT FROM ST. NO.,2I4,F10.2)
    5 FORMAT(* WAVENUMBER  WAVELENGTH  TRANS 1  ABSORB 1  TRANS 2  ABSOR
     1B 2  TRANS 3  ABSORB 3  TRANS 4  ABSORB 4  TRANS 5  ABSORB 5*/)
    6 FORMAT(F8.2,5F10.5)
    7 FORMAT(11H1PRESSURE =F9.5,3X13HTEMPERATURE =F9.2,3X8HSLIT A =F9.5,
     13X4HDV =F7.4,3X6HDELV =F7.4,3X7HBOUND =F7.2/)
    8 FORMAT(F12.4,10X,E11.4)
    9 FORMAT(///33H0DATA FOR PLOTTING ARE ON TAPE C-,I4)
   10 FORMAT(I4)
   11 FORMAT(60H1TOO MANY LINE CARDS,EXCEEDS DIMENSION. LAST CARD READ W
     1AS =F9.2)
   12 FORMAT(19H1I7 TOO LARGE. I7 =I5,5X9HGNU(I7) =F9.2,5X12HV1-A+BOUND
     1=F9.2)
   13 READ 1,(FMT1(I),I=1,9)
   14 READ 1,(FMT2(I),I=1,9)
   15 READ 1,(FMT3(I),I=1,9)
   16 READ 1,(FMT4(I),I = 1,9)
      PRINT1,FMT1,FMT2,FMT3,FMT4
   17 READ FMT3,K1,I7,DV,DELV,V1,V2,A,P,NP,BOUND
   18 READ FMT4,TEMP,TEMPO,BX,CX
 1800 READ FMT2,(W(I),I=1,K1)
C
C     SET UP TEMPERATURE CORRECTION CONSTANTS.
C
 1810 CS1 = (TEMPO-TEMP)/(TEMPO*TEMP*.6950)
 1820 CS2=(TEMPO/TEMP)**BX
 1830 CA=((TEMPO/TEMP)**CX)*P
      CALL Q9 EXUN
C     READ IN LINE DATA AND COUNT CARDS
C
      VLO=V1-A-BOUND-1.$      VHI=V2+A+BOUND
   20 I=0
   21 I=I+1
C
C
   25 READ   FMT1 ,GN ,S,ALF,EPP,MOL
      IF(GN .LT.VLO) GO TO 25
      IF(GN .GT.VHI) GO TO 27
      IF(EOF,60)27,272
  272 IF(MOL.EQ.1)GO TO 274
  273 ALF=.1
  274 AA(I)=(ALF*CA)**2
      GNU(I)=GN
 2501 SA(I)=S*ALF*CA*CS2*EXP(-EPP*CS1)            $ GO TO 21
   27 I1=I-1
   28 FORMAT(2F10.3,I5)
      PRINT 28, GNU(1),GNU(I1),I1
```

```
C
C CHECK THAT GNU(I7) IS NOT GREATER THAN V1-A+BOUND.
C
   32 XBOUND = V1 -A + BOUND
   33 IF(GNU(I7) .GT. XBOUND)34,36
   34 PRINT 12,I7,GNU(I7),XBOUND
   35 GO TO 1001
   36 PRINT 7,P,TEMP,A,DV,DELV,BOUND
   37 PRINT 38,(W(K),K=1,K1)
   38 FORMAT(26X,*W1 =*E10.3,5X,*W2 =*E10.3,5X*W3 =*E10.3,5X *W4 =*E10.3
     1,5X*W5 =*E10.3/)
   39 PRINT 5
C
C      SET UP CALCULATION COUNTERS.
C
   40 I4=1
   41 SLIT=2.*A
   46 IDIV=DELV/DV+.01
   47 NODD1=IDIV+1
   48 NODD2 = (SLIT/DV) + 1.00001
   49 NODD3 = NODD2 - IDIV
   52 NJ = 1
   55 SLTFTR = DV/(A**2)
   68 VI = V1
   69 V = V1 - A
C
C      INITIALIZE SUMMATIONS
C
   70 DO 71 K=1,K1
   71 SUMT(K) = TRAPSUM(K) = 0.0
C
C      ABSORPTION COEFFICIENT CALCULATION BY SUBROUTINE ABSCOEF
C
   79 DO 150 NN = NJ,NODD2
   80 CALL ABSCOEF(V,BOUND,I1,P)
C
C      TRANSMISSION SUMMATION AND APPLICATION OF VARIABLE PORTION OF SLIT
C      FACTOR INTEGRAL
C
  130 DO 138 K=1,K1
  134 T(NN,K) = EXPF(-W(K)*CAY)
  136 TT(K)=(A-ABSF(V-VI))*T(NN,K)
  138 SUMT(K)=SUMT(K)+TT(K)
  150 V=V+DV
  154 V=VI
C
C      TRANSMISSION AND ABSORPTION CALCULATION WITH FINAL SLIT FACTOR
C      APPLICATION
C
  160 DO 175 K=1,K1
  161 TF(K)=SUMT(K)*SLTFTR
  162 E(K)=1.-TF(K)
  163 IF(V-V1) 509,164,166
  164 TRAPSUM(K) = TRAPSUM(K) +0.5*E(K)
  165 GO TO 174
  166 TRAPSUM(K) = TRAPSUM(K) + E(K)
C
C      OUTPUT STATEMENTS AND FREQUENCY INCREMENTING
C
  174 ALAM = 1.E+4/V
  175 CONTINUE
  176 PRINT2, V,ALAM,(TF(K),E(K),K=1,K1)
      PUNCH 6,V,(TF(K),E(K),K=1,K1)
  177 GO TO(180,181),NP
  180 WRITE(2, 6  )V,(TF(K),K=1,K1)
  181 V=V+DELV
  190 DO 191 K=1,K1
  191 SUMT(K)=0.0
  200 IF(V-V2)201,201,235
C
C      SAVING OF ALREADY CALCULATED TRANSMISSION VALUES FOR NEXT STEP
C      IN THE CALCULATION
C
  201 DO 205 K=1,K1
  202 DO 204 NN = NODD1,NODD2
  203 KK=NN-IDIV
  204 T(KK,K)=T(NN,K)
  205 CONTINUE
  210 VI=V
  211 V=V-A
  215 DO 219 NN = 1,NODD3
  216 DO 218 K=1,K1
  217 TT(K)=(A-ABSF(V-VI))*T(NN,K)
  218 SUMT(K)=SUMT(K)+TT(K)
  219 V=V+DV
  225 NJ = NODD3 + 1
  229 GO TO 79
C
C      TRAPAZOIDAL RULE CALCULATION OF TOTAL ABSORPTANCE
C
  235 PRINT 3,P,A,TEMP
  236 DO 238 K=1,K1
  237 AB(K) = DELV*(TRAPSUM(K)-0.5*E(K))
  238 PRINT 8,AB(K),W(K)
  239 READ 240,V1,V2
  240 FORMAT(2F10.2)
  241 IF(EOF,60) 250,68
  250 GO TO(251,1001),NP
  251 END FILE 2
  252 REWIND 2
  255 GO TO 1001
C
C      KICKOUT STATEMENTS
C
  500 NK = 31
  501 GO TO 1000
  509 NK = 163
 1000 PRINT 4,NK,I1,GNU(I1)
 1001 CALL EXIT
      END
      SUBROUTINE ABSCOEF(V,BOUND,I1,P)
      COMMON/1/GNU(3000),SA(3000),AA (3000),CAY(1)/2/T(3000,5),I5
C
C     DETERMINATION OF INDEXING VALUES FOR ABSORPTION COEFFICIENT CALC.
C
   10 DO 14 I =I5,I1
   11 IF(V-BOUND-GNU(I))12,12,14
   12 I5 = I
   13 GO TO 15
   14 CONTINUE
   15 DO 19 K = I5,I1
   16 IF(V+BOUND-GNU(K))17,17,19
   17 I6 = K
   18 GO TO 25
   19 CONTINUE
C
C     CALCULATION OF ABSORPTION COEFFICIENT WITH LORENTZ LINE SHAPE
C
   25 CAY1 = CAY2 = 0.0
   30 DO 46 I = I5,I6
   32 Y = ABSF(V-GNU(I))
   34 IF(Y-2.)36,36,42
   36 SUM1=SA(I)/(Y**2+AA(I))
   38 CAY1 = CAY1 + SUM1
   40 GO TO 46
   42 SUM2=SA(I)/Y**2
   44 CAY2 = CAY2 + SUM2
   46 CONTINUE
   50 CAY = 0.3183*(CAY1+CAY2)
   52 RETURN
      END
         SCOPE
(LOAD
(RUN, 2,10000
 (F10.3,E10.3,F5.3,F10.3,43X,I2)
 (5E10.3)
 (I2,I4,2F6.3,2F8.2,F5.2,E10.3,I2,F7.2)
```

```
         SUBROUTINE ABSCOEF(V,BOUND,I1,P)
         COMMON/1/GNU(3000),SA(3000),AA (3000),CAY(1)/2/T(3000,5),15
c
c        DETERMINATION OF INDEXING VALUES FOR ABSORPTION COEFFICIENT CALC.
c
    10   DO 14 I =15.I1
    11   IF(V-BOUND-GNU(I))12,12,14
    12   15 = I
    13   GO TO 15
    14   CONTINUE
    15   DO 19 K = 15,I1
    16   IF(V+BOUND-GNU(K))17,17,19
    17   I6 = K
    18   GO TO 25
    19   CONTINUE
c
c        CALCULATION OF ABSORPTION COEFFICIENT WITH LORENTZ LINE SHAPE
c
    25   CAY1 = CAY2 = 0.0
    30   DO 46 I = 15,I6
    32   Y = ABSF(V-GNU(I))
    34   IF(Y-2.)36,36,42
    36   SUM1=SA(I)/(Y**2+AA(I))
    38   CAY1 = CAY1 + SUM1
    40   GO TO 46
    42   SUM2=SA(I)/Y**2
    44   CAY2 = CAY2 + SUM2
    46   CONTINUE
    50   CAY = 0.3183*(CAY1+CAY2)
    52   RETURN
         END

SINGLE-BANK COMPILATION.
```

FORTRAN PROGRAM
CDC 3800

MORTRAN
Dana Gregory
June 29, 1973

Given the original gas concentration, W_o, the corresponding theoretical transmission, $T(W_o, \nu_j)$, as a function of the wavenumber ν_j, and the original gas concentration, W_o, MORTRAN will compute theoretical transmissions, $T1(W_i,\nu_j)$, for the desired concentrations W_i.

Used in the computation is Beer's Formula,

$$T1(W_i,\nu_j) = T(W_o,\nu_j)^{W_i/W_o},$$

where $T1(W_i\nu_j)$ is a function of the set of transmission values for each desired concentration.

Input

Card A

Column	Field	Description
1-10	F10.2	Beginning wavenumber
11-20	F10.2	Ending wavenumber
21-30	F10.2	Wavenumber increment
31-40	E10.3	Original concentration, W_o
46-50	I5	NW, number of desired concentrations to be used in signal run

Card B

1-10	E10.3	1st desired concentration, W_1
11-20	E10.3	2d desired concentration, W_2

MORTRAN - (cont.)

Column	Field	Description
.	.	.
.	.	.
71-80	E10.3	i^{th} desired concentration, W_i
Card C_i-C_n		
1-10	F10.0	Wavenumber, ν_j
11-20	F10.0	Transmission, $T(W_o,\nu_j)$

Output

Card A_i		
1-10	F10.2	Beginning wavenumber
11-20	F10.2	Ending wavenumber
21-30	F10.2	Wavenumber increment
31-40	E10.3	Concentration, W_i $i = 0; i = 1, NW$
Card B_i-B_n		
1-5	I5	Sequence number
6-10	3PF5.0	Transmission, $T(W_i,\nu_j)$
11-15	3PF5.0	Transmission, $T(W_i,\nu_{j+1})$
.	.	.
.	.	.
.	.	.
76-80	3PF5.0	Transmission, $T(W_i,\nu_{j+14})$

SUBROUTINES USED: None.

MORTRAN: Computes additional transmissions and outputs in packed format.

I

```
      PROGRAM MORTRAN
      DIMENSION WN(10),T(5000),V(5000),T1(5000)
01    READ 02, BV,EV,DV,W,NW
      IF(EOF,60)850,03
02    FORMAT(3F10.2,E10.3,5X,I5)
03    NT=1.00001+(EV-BV)/DV
      NWA=XABSF(NW)
      IF(NT.LT.5000)04,030
04    READ 05, (WN(I),I=1,NWA)
05    FORMAT(5E10.3)
      DO 09 I=1,NT
      READ 06, V(I),T(I),WT
06    FORMAT(F10.2,E10.3,40X,E10.3)
      IF(NW.LE.0)07,065
065   IF(W.EQ.WT)07,910
07    IF(V(I).EQ.BV)09,08
08    X=V(I-1)+DV
      Y=V(I)
      IF(ABS(Y-X).LE..0001)09,410
09    CONTINUE
      NCARDS=NT/15
      IF(NCARDS*15.NE.NT)21,22
21    NCARDS=NCARDS+1
22    IB=1
      IE=IB+14
      PRINT 3001,BV,EV,DV,W
      PUNCH 20,BV,EV,DV,W
      DO 259 K=1,NCARDS
      PUNCH 30,K,(T(M),M=IB,IE)
      PRINT 3002,K,(T(M),M=IB,IE)
      IB=IB+15
      IE=XMINOF(IE+15,NT)
259   CONTINUE
      DO 40 J=1,NWA
      EXPT=WN(J)/W
      DO 10 I=1,NT
10    T1(I)=T(I)**EXPT
      PUNCH 20, BV,EV,DV,WN(J)
20    FORMAT(3F10.2,E10.3)
      PRINT 3001, BV,EV,DV,WN(J)
3001  FORMAT(35X,3F10.2,E10.3)
      IB=1
      IE=IB+14
      DO 40 K=1,NCARDS
      PUNCH 30, K,(T1(M),M=IB,IE)
30    FORMAT(I5,(15(3PF5.0)))
      PRINT 3002,K,(T1(L),L=IB,IE)
3002  FORMAT(2X,I5,(15(3PF5.0)))
      IB=IB+15
      IE=IE+15
      IE=XMINOF(IE,NT)
40    CONTINUE
      GO TO 01
850   PRINT 851
851   FORMAT(12(/),30X,*SORRY, THERE ARE NO MORE CARDS*)
```

Reproduced from best available copy.

```
I         MOPTRAN

      CALL EXIT
 900  PRINT 901, NT,RV,EV,DV
 901  FORMAT(20X*ERROR IN CARDS*//7X*NT=*,I5,*RV=*,F10.2,2X*EV=*,F10.2,
     $ 2X,*DV=*,F10.2)
      CALL EXIT
 910  PRINT 911, W,WT,Y,X
 911  FORMAT(10X*ERROR IN DATA*//10X,2E10.3/10X,2F10.3)
      CALL EXIT
      END

 SINGLE-BANK COMPILATION.
```

APPENDIX 2

TABLE OF CONTENTS

Page

A2. Program EPAGAS Documentation and Analysis 111

A2.1 Program Summary 111

A2.1.1 List of Fortran Symbols and Descriptions 112

A2.1.2 List of Program and Subprogram Descriptions 121

EPAGAS
CALC
CLOSER
CONST
CXMIZ
EROR1
EROR2
FINDIT
FINDTRN
GAUSSL
GRAPHS
IBINSER
MINMYZD
PARTIN
PLOTD
PLOTT
PRINT1
PRINT2
PRINT3
READALL
READ1
READ2
READ3
READ4
READ6
READ7
READ8
RMSERR
THEOTRN

A2.2 Simulation of Gaussian Noise 170

A2.3 Constituent and Transmission Error Formulation 171

A2.4 Interpolation Formulas Used in Transmission Library 173

A2.5 Program Execution Techniques 177

A2.5.1 Theoretical Gas Library Spectra 177

A2.5.2 Experimental Transmission Data 178

A2.5.3 Minimization of "Goodness of Fit" Controls 179

A2.5.4 Error Parameters 180

A2.5.5 Miscellaneous Execution Techniques 182

APPENDIX 2

A2. PROGRAM EPAGAS DOCUMENTATION ANALYSIS

A2.1 PROGRAM SUMMARY

In this appendix, the main minimization program EPAGAS and all of the subprograms needed by EPAGAS are described. All coding was done in the FORTRAN applicable to the CDC 3800 computer. Each description lists what that particular subprogram does, what input parameters are needed, and what output parameters are produced.

The computer program must be given: (1) a library of theoretical transmissions for each gas constituent which might possibly be present in the measured total transmissions. For each constituent, the theoretical transmissions are known functions of both wavenumber and gas concentration. The library must be large enough to bound the regions of interest in both wavenumber and concentrations to be expected. See THEOTRN description for exact preparation of the library table. Due to possible computer memory limitation, the user may wish to "tailor-make" separate libraries for different experimental situations. However, care must be taken in selecting proper wavenumber spacings and individual concentrations so that double linear interpolation within the transmission table will result in correct transmissions values for wavenumbers and concentrations lying between grid points of the library table.

(2) Next, a set of measured experimental total transmissions TG as a function only of known wavenumber ν must be provided. This spectrum, obtained from the spectrophotometer, will be compared to a trial spectrum TC which is a function of gas constituent, gas concentration W, and wavenumber ν as determined from the library of theoretical transmissions. By adjusting the amounts of W for each gas constituent as a function of wavenumber, the program tries to get TC to agree as best as possible with TG, thus determining the gas composition and the amounts of each gas constituent that were present in the sample producing the total transmission, TG. READ7 and/or READ8 descriptions provide specific information on the preparation details for inputting TG.

(3) The program must also be provided with a starting set of gas concentrations W for each gas constituent that is suspected to be present in the sample producing the given experimental total transmission TG. Also given is a corresponding set of concentration increments, WI, which are used when the minimization process tries to vary the individual W_i by an amount WI_i to find the best agreement between TG and TC. See the READ6 description for preparation of W and WI.

(4) Along with the initial values of W and WI, the program also needs error bounds for each of the individual gas constituent concentrations. See EROR1 description for detail error formulas. The errors defined on the gas concentrations are used primarily to help guarantee an acceptable solution by keeping the gas concentrations within physically reasonable bounds. READ2 description provides details for inputting the error parameters.

(5) Finally, various program controls must be provided. Some of these control parameters are used in directing the minimization process and are listed in detail in the READ3 description. The remaining control parameters are used in directing printing and plotting of intermediate results and in providing certain convergence criterion. Details for these parameters are found in the READ4 description.

A2.1.1 List of Fortran Symbols and Descriptions

Single-Valued Variables

Variable	Description
CFM	A constant confidence factor to be applied to the experimental transmission data if individual ones are not specified.
GLB	The minimum root-mean-square (rms) error permitted before using the minimization routine MINMYZD again.

Variable	Description
IPLOT	A plotting control which, when set to m, permits a graph of the given and currently computed transmissions, as functions of wavenumber, to be plotted after every m^{th} use of MINMYZD.
IPRINT	A printing control which, when set to n, permits printing of intermediate results after every n^{th} use of MINMYZD.
MK	The maximum number of uses of MINMYZD permitted for a particular data set.
N	N = NG + NT. The total number of errors in the error array ERR.
NG	The total number of gas constituents to be found in the library of theoretical transmissions.
NT	The number of experimental transmission values used.
Q	The minimum percent by which each try of MINMYZD must improve the mean square deviation.
RMS	Current root-mean-square (rms) error computed from the error array ERR.
RMSTP	Lowest root-mean-square (rms) error computed from the error array ERR.
RMST	Lowest root-mean-square (rms) error computed from the NG+1 through NG+NT elements of the error array ERR. This rms value corresponds to the errors based on the difference between the given experimental transmission values and the theoretical transmission values currently computed.
RMSW	Lowest root-mean-square (rms) error computed from the 1 through NG elements of the error array ERR. This rms value corresponds to the errors defined on the individual gas concentrations.

Variable	Description
UB	Beginning wavenumber corresponding to the given transmission values if they are equally spaced in wavenumber.
UE	End wavenumber corresponding to the given transmission values if they are equally spaced in wavenumber.
UI	Wavenumber increment corresponding to the given transmission values if they are equally spaced in wavenumber.
UMAX	Maximum wavenumber value for which the given transmission values will range over.
UMIN	Minimum wavenumber value for which the given transmission values will range over.
ZEP	A number used by MINMYZD which tells how much of the predicted final step to take.

Array Variables

Variable	Maximum dimension	Description
CF	150	Confidence factor, $0 \leq CF(K)$, K=1, NT, corresponding to the NT given transmission values.
ERR	160	Errors based on the individual gas concentrations and the weighted differences between the given and computed transmission values. Elements 1 through NG are the errors corresponding to the gas concentrations, and elements NG+1 through NG+NT are the errors corresponding to the transmission errors.

Variable	Maximum dimension	Description
IDGAS	10	Integer numerical identifiers (I2) for each of the gases to be found in the gas library of theoretical transmissions. IDGAS(K), K=1, NG correspond to the K^{th} gas.
IPW	10	One of the error parameters used in computing the errors based on the individual gas concentrations. This parameter is an odd positive integer used as a power in the error function. IPW controls the rate of increase of the error as the current gas concentration ventures outside the acceptable range for that particular concentration. IPW(K), K=1, NG correspond to the K^{th} gas.
KX	10	The first six elements of the array KX are input control parameters used by the main minimization routine MINMYZD. The last four elements of the array are output parameters of MINMYZD and provide various convergence information. KX(1) is the maximum number of new sets of errors (minimization cycles) to use. If KX(1)=0, then one cycle is used. If KX(1)<0, then each gas concentration increment WI is reduced by 10 percent after each complete use of MINMYZD, and $\lvert KX(1) \rvert$ is used for the maximum number of cycles to be used in one complete use of MINMYZD.

Variable	Maximum dimension	Description
		KX(2) is the number of cycles to count for a bad cycle, that is, one which did not result in enough improvement, $KX(2) \geq 1$.
		KX(3) is the percent improvement in tenths in the rms error for the cycle to be counted as a good cycle. Any cycle resulting in less than this causes KX(2) cycles to be counted rather than one cycle, $KX(3) \geq 0$.
		KX(4) is the limit as to the multiple of a WI increment which may be taken in one full step (ZEP=1). If KX(4)=0, then 100 is used. If KX<0, then no step will be permitted in the -WI direction for the first complete use of MINMYZD. Future uses of MINMYZD uses $\vert KX(4) \vert$.
		KX(5) is the minimum ZEP in 0.001's permitted. KX>0.
		KX(6) is the minimum total percent improvement in tenths for ZEP not to be changed. KX(6)>0.
		KX(7): If KX(7)=1, then at least one single step (even if it were a step of WI_i) resulted in an improvement of at least KX(3); otherwise, KX(7)=0.
		KX(8): If KX(8)=1, then there was some improvement sometime; otherwise, KX(8)=0.

Variable	Maximum dimension	Description
		KX(9): If KX(9)=-1, then all WI's were too small. Convergence has probably been reached. If KX(9)=0, then the solution matrix was singular. If KX(9)=1, then the matrix was normal.
		KX(10) is the actual total percent improvement in tenths.
NU	10	Number of wavenumber values used for each gas in the library of theoretical transmissions. NU(K) is the number of wavenumber values used for the K^{th} gas, K=1, NG.
NW	10	Number of concentration values used for each gas in the library of theoretical transmissions. NW(K) is the number of concentrations used for the K^{th} gas, K=1, NG.
S	5	An array used by READALL and its associated reading routines. If S(K)=0, then the input information was correct. If S(K)≠0, then a reading error occurred.
TC	150	Computed theoretical transmission values, TC(K), K=1, NT based on the NG current gas concentrations being used. TC(K) corresponds to the K^{th} value. $TC(K) = \prod_{I=1}^{NG} TCI(I,K)$.
TCI	10	Individual computed theoretical transmission values TCI(I,K) for a given wavenumber U(K), K=1, NT and given gas concentrations W(I), I=1, NG.

Variable	Maximum dimension	Description
TG	150	Given experimental transmission values TG(K), K=1, NT as a function of known wavenumber U(K), K=1, NT.
TRAN	(5000,5)	A two-dimensional array which contains all of the theoretical transmission values that make up the gas library. The transmission values are a function of known wavenumbers, known gas constituents, and known concentrations for each of the specified gas constituents. See THEOTRN description for exact formulation of the array.
U	150	Known wavenumbers, U(K), K=1, NT corresponding to the array TG of the NT given experimental transmission values.
UR	(10,3)	A two-dimensional array UR(I,J) containing the wavenumber range and spacing used for each of the gas constituents in the gas library. I corresponds to the Ith gas. J=1 corresponds to the beginning wavenumber, J=2 corresponds to the end wavenumber, and J=3 corresponds to the wavenumber increment.
WHR	10	One of the error parameters used in computing the errors based on the individual gas concentrations. This parameter is the half-range of permitted variation in concentration for the Kth gas constituent and is

Variable	Maximum dimension	Description
		simply the difference between the mean concentration WM(K) and the lowest concentration of the Kth gas constituent.
WIO	10	Original increments to be applied to the starting gas concentration WO(K), K=1, NG. These are needed by MINMYZD when it tries to adjust the original concentrations WO to get the best agreement between the given and computed transmission values. These values may be modified in the minimization process, and hence become WI(K), K=1, NG.
WI	10	Current incremental values of gas concentrations, WI(K), K=1, NG.
WM	10	One of the error parameters used in computing the errors based on the individual gas concentrations. This parameter is the mean concentration value of the desired upper and lower bounds of the concentrations of the Kth gas constituent, K=1, NG.
WO	10	Original gas concentrations to be used in the minimization process. WO(K) corresponds to the Kth gas constituent, K=1, NG.

Variable	Maximum dimension	Description
WR	(10,5)	A two-dimensional array WR(I,J) containing the actual gas concentrations used in the gas library. I corresponds to the I^{th} gas constituent, and J corresponds to the J^{th} gas concentration of the I^{th} gas constituent.
W	10	Current gas concentrations being used in the minimization process. W(K), K=1, NG corresponds to the K^{th} gas constituent.
WW	10	One of the error parameters used in computing the errors based on the individual gas concentrations. This parameter is a "normalization" factor and is used to scale the gas concentration errors. WW(K), K=1, NG are the factor for the K^{th} gas constituent.

A.2.1.2 List of Program and Subprogram Descriptions

FORTRAN PROGRAM
CDC 3800

EPAGAS
Margot Ackley
July 30, 1973

Program EPAGAS is a computer program which is designed to determine atmospheric composition from measured total transmissions at known wavenumbers. The program utilizes a set of theoretical transmission spectra of selected gas constituents at known values of concentrations and at known wavenumbers, and a set of observed total transmissions at known wavenumbers only. By employing a particular minimization technique, EPAGAS tries to fit the observed transmissions to the theoretical transmissions by varying the concentrations of the individual constituents.

Assuming a certain concentration of each gas constituent in an atmospheric path, a corresponding transmission can be found from a large table of known transmissions as a function of gas constituent, wavenumber, and gas concentration. The total transmission at a specified wavenumber is then computed and compared to the observed transmission value obtained from a spectrophotometer. The difference between the computed and observed transmissions is then minimized by varying the concentrations of the individual gas constituents. The minimization process tries to reduce the rms error of an error function defined on the errors of the given and computed transmissions and on the concentrations. (See A2.3.) The concentration errors are used mainly to help keep the concentrations within such a range as to provide physically realistic solutions.

The program uses various controls to direct its progress in the minimization process. The initial and final gas constituent concentrations, computed transmissions, and errors are always printed together with a plot of the given and computed transmissions as a function of wavenumber. Intermediate results may also be printed and plotted.

For preparation of the theoretical gas library table, see the THEOTRN description. For preparation of the given transmission data and all other input needed by the program, see the READALL description. It is imperative that all descriptions pertaining to the program be read and understood.

USES:

- CALC
- CLOSER
- CONST
- CXMIZ
- EROR1
- EROR2
- FINDIT
- FINDTRN
- GAUSSL
- GRAPHS
- IBINSER
- MINMYZD
- PARTIN
- PLOTD
- PLOTT
- PRINT1
- PRINT2
- PRINT3
- READALL
- READ1
- READ2
- READ3
- READ4
- READ6
- READ7
- READ8
- RMSERR
- THEOTRN

FORTRAN SUBROUTINE
CDC 3800
(Internal to EPAGAS)
CALC (W, ERR, RMS)

CALC
Margot Ackley
July 30, 1973

Given the array W of NG current gas concentrations and the selected arguments from the two COMMON blocks, COMMON/100/ and COMMON/101/, CALC computes an error array ERR of length NG + NT, where NT = the number of computed and given transmissions and NG = the number of gas concentrations. CALC also computes the rms error RMS needed by MINMYZD, and the individual rms errors RMSW for the error based only on the concentrations, and RMST for the error based only on the transmissions. (See A2.3.)

The two COMMON blocks utilized are as follows: COMMON/100/TRAN (5000,5), NU(10), NW(10), UR(10,3), WR(10,5), TG(150), U(150), CF(150), TC(150), WM(10), WHR(10), WW(10), IPW(10); COMMON/101/NG, NT, RMSTP, RMST, RMSW.

The first NG elements of the array ERR are the errors based on errors from the NG gas concentrations. (See EROR1 description.) The NG + 1 elements through the NG + NT elements of the array ERR are the errors based on errors from the NT given and currently computed transmissions. (See EROR2 description.)

Before computing the individual transmission errors, CALC must first calculate the NT computed transmissions TC based on the current NG gas concentrations W given to CALC. CALC uses subroutine FINDTRN to find the individual computed transmission TCI(i,j) for a given wavenumber U(j) and a given gas concentration W(i). The total transmission for the particular wavenumber U(j) is then the product of all the transmissions for all the gas concentrations W(i) for the particular U(j) where:

$$TC(j) = \prod_{i=1}^{NG} TCI(i,j) \quad .$$

If the current use of CALC produces an RMS error which is smaller than the previous rms error, then RMSTP = RMS, and the array TC will then contain the computed transmissions producing the lower RMS. If RMS is greater, then the values RMSTP and TC will contain the last set of values resulting in an improved rms error.

For definitions of arguments and variables used, see below descriptions:

EPAGAS: W
READ7(8): NT, U, TG, CF
THEOTRN: NG

USES: EROR1, EROR2, FINDTRN(IBINSER), and RMSERR.

CALC: Computes concentration and transmission errors for MINMYZD.

FORTRAN FUNCTION
CDC 3800
(Internal to EPAGAS)
CLOSER (A, X, B)

CLOSER
J. R. Winkelman
June 15, 1965

CLOSER is the value of X in the closed region $A \leq X \leq B$, taking on the values of the end points for X outside the range. More exactly, if $X \leq A$, the value of CLOSER will be A. If $X \geq B$, the value of CLOSER will be B. Otherwise, CLOSER=X.

USES: No subroutines.
CLOSER: Restrict X to a closed region.

FORTRAN SUBROUTINE
CDC 3800
(Internal to EPAGAS)

CONST.
Margot Ackley
July 30, 1973

CONST (S, UMIN, UMAX, ZEP, KX, Q, MK, GLB, CFM, IPRINT, IPLOT)

CONST sets certain constants needed by program EPAGAS and its associated subroutines. These values are as follows:

S(1) = S(2) = -1.
S(3) = S(4) = S(5) = 0.

WM(1) = 3.237E22
WM(2) = 1.235E20
WM(3) = 1.235E20
WM(4) = 1.235E20
WM(5) = 1.245E20
WM(6) = 0.715E20

WHR(1) = 3.437E22
WHR(2) = 1.235E20
WHR(3) = 1.235E20
WHR(4) = 1.235E20
WHR(5) = 1.245E20
WHR(6) = 0.715E20

WW(1) = WW(2) = WW(3) = WW(4) = WW(5) = WW(6) = 1
IPW(1) = IPW(2) = IPW(3) = IPW(4) = IPW(5) = IPW(6) = 1

UMIN = 1000.
UMAX = 1250.

KX(1) = 10
KX(2) = 3
KX(3) = 50
KX(4) = 30
KX(5) = 100
KX(6) = 100
KX(7) = 0
KX(8) = KX(9) = KX(10) = 0

ZEP = 0.6
Q = 10.0
GLB = 3.0
CFM = 10.0
MK = 10
IPRINT = 1
IPLOT = 1

The definitions of the above arguments are given in the below subroutine descriptions.

S	READALL
WM, WHR, WW, IPW	READ2
UMIN, UMAX	READ1
ZEP, KX	READ3
Q, GLB, CFM, MK IPRINT, IPLOT	READ4

CONST utilizes COMMON/100/ where: COMMON/100/TRAN(5000,5), NU(10), NW(10), UR(10,3), WR(10,5), TG(150), U(150), CF(150), TC(150), WM(10), WHR(10), WW(10), IPW(10).

USES: No subroutines.
CONST: Set values for EPAGAS.

FORTRAN FUNCTION
CDC 3800
(Internal to EPAGAS)
CXMIZ (KX, G, GLB, MK, RMS)

CXMIZ
Margot Ackley
July 30, 1973

Given the array KX (see READ3 description), the arguments Q and GLB (see READ4 description), MK the remaining uses of MINMYZD to be permitted (see READ4 description), and RMS, the current rms error, CXMIZ checks to see if MINMYZD is to be used again.

If KX(9) = -1, then CXMIZ = 2.0 and returns. If KX(9) = 0, then CXMIZ = 3.0 and returns. If KX(9) = 1, then CXMIZ checks the following: if $MK \leq 0$, if $\frac{Q}{.1} > KX(10)$, or if $RMS \leq GLB$, then CXMIZ = 2.0.

If none of the above occur, then CXMIZ = 1.0 and returns.

Where:

CXMIZ = 1, use MINMYZD again,
CXMIZ = 2, do not use MINMYZD again,
CXMIZ = 3, singular matrix, halt computation.

USES: No subroutines.
CXMIZ: Checks for future uses of MINMYZD in EPAGAS.

FORTRAN SUBROUTINE
CDC 3800
(Internal to EPAGAS)
EROR1 (W, ERR)

EROR1
Margot Ackley
July 30, 1973

Given NG values of W, where W(I) is the concentration of the I^{th} gas constituent and corresponding error parameters WM, WHR, WW, and IPW, EROR1 computes the NG errors ERR as defined below:

$$ERR(I) = WW(I)^* \left(\frac{W(I)-WM(I)}{WHR(I)} \right)^{IPW(I)} .$$

$$I = 1, NG$$

WM, WHR, WW, and IPW are provided through COMMON/100/: COMMON/100/ TRAN(5000,5), NU(10), NW(10), UR(10,3), WR(10,5), TG(150), U(150), CF(150), TC(150), WM(10), WHR(10), WW(10), IPW(10).

USES: No subroutines.
EROR1: Computes gas constituent errors for EPAGAS.

FORTRAN FUNCTION
CDC 3800
(Internal to EPAGAS)
EROR2 (TG, TC, CF)

EROR2
Margot Ackley
July 30, 1973

EROR2 is the error computed from the discrepancy between the given transmission TG and the computed transmission TC, weighted by a confidence factor CF, where

EROR2 = CF*(TG-TC).

USES: No subroutines.
EROR2: Computes transmission errors for EPAGAS.

FORTRAN FUNCTION
CDC 3800
(Internal to EPAGAS)
FINDIT (A, L, M, N)

FINDIT

J. R. Winkelman
December 18, 1962

The value of FINDIT will be that A(L) which would be obtained if the array A was sorted between A(M) and A(N). A(L) has its correct new value, that is, $A(K) \leq A(L)$ for $M \leq K \leq L$, $A(K) \geq A(L)$ for $L \leq K \leq N$. Thus, the array is in better sort on exit from FINDIT. The array A outside of the range M, N is unchanged.

Example: To find the 20 percentile and median of the array from A(10) to A(20),

```
PMED = FINDIT (A, 15, 10, 20)
P20 = FINDIT (A, 12, 10, 14) or FINDIT (A, 12, 10, 20).
```

USES: PARTIN.

FINDIT: Find the Lth sorted value.

FORTRAN SUBROUTINE
CDC 3800
(Internal to EPAGAS)
FINDTRN (W, UP, TT)

FINDTRN
Margot Ackley
July 30, 1973

Given the arguments W and UP and selected variables from the two COMMON blocks--COMMON/100/TRAN(5000,5), NU(10), NW(10), UR(10,3), WR(10,5), TG(150), U(150), CF(150), TC(150), WM(10), WHR(10), WW(10), IPW(10); and COMMON/101/NG, NT, RMSTP, RMST, RMSW--FINDTRN finds the corresponding theoretical transmissions TT.

The array W contains the NG current values of gas concentrations, and UP is some given wavenumber within the range of interest. The computed theoretical transmissions are placed in the array TT, which has a maximum length of 10, where TT(I) corresponds to the theoretical transmissions from the gas library as a function of wavenumber UP and concentration W(I) of the I^{th} gas.

Because the current concentrations W and wavenumber UP may not necessarily correspond to those wavenumbers and concentrations in the gas library table TRAN (see THEOTRN description), FINDTRN uses double linear interpolation within the two-dimensional table TRAN to determine the transmission TT(I) corresponding to W(I) and UP.

Wavenumber UP will never lie outside the range of the table. However, the concentrations W could possibly lie outside the range due to the minimization process which varies only the array W. In the event that the individual W(I) lies above the range of the concentrations of the I^{th} gas, the highest two concentrations of that gas are used to determine a transmission value TT(I) for W(I) and UP. In the case where W(I) lies below the range of the table, the smallest concentration in the table and a concentration of zero are used in determining the corresponding transmission. The interpolation formula for this case utilizes the fact that, by definition, a transmission value for zero concentration is 1. (See A2.4.) The interpolation process can possibly result in transmissions outside the range 0 to 1, but the structure of the error functions tends to discourage selections of those concentrations W which make this occur.

USES: IBINSER.

FINDTRN: Finds theoretical transmissions as a function of wavenumber and concentrations for EPAGAS.

FORTRAN SUBROUTINE
CDC 3800
(Internal to EPAGAS)
GAUSSL (A, NRD, NR, NC, D)

GAUSSL
L. David Lewis
January 20, 1970

Given an augmented matrix A, its row dimension NRD (from the DIMENSION statement defining A), the number of rows NR (equals the order of the matrix), and the number of columns NC (equals NR plus the number of vectors);

Find D, the determinant of the matrix and the solution vectors of the indicated linear system. The given matrix is destroyed, and the given vectors are replaced by the solution vectors.

Restrictions: A must be stored columnwise;
$NC \geq NR$, $NR \leq NRD$. If NC = NR, only D is computed.

Method: Gauss elimination with pivotal condensation.

Error conditions: D = 0, the matrix was singular; the augmented matrix has been destroyed.

USES: No subroutines.
GAUSSL: Determinant of a matrix by Gauss elimination and solution.

FORTRAN SUBROUTINE **GRAPHS**
CDC 3800 **Margot Ackley**
(Internal to EPAGAS) **July 30, 1973**
GRAPHS (A, B, Y, NGC, NXC, NUC, NŌC, NBC, NQ)

Edits an array NQ of 101 alphanumeric characters (R1 format) to allow computer plotting of functions.

For A<B, A is the minimum ordinate associated with NQ(1) and B is the maximum ordinate associated with NQ(101). Y is the value to be plotted. If $A \leq Y \leq B$, the quantity D = 1 + 100*(Y-A)/(B-A) is computed and rounded to the nearest integer I. Then NQ is examined. If NQ(1) contains the blank character NBC or a decimal point (.), then the graph character NGC is stored there; if NQ(I) does not contain NBC or a decimal point, then the intersection character NXC is stored in NQ(I). If Y is outside the range A,B , then I is computed as though Y were in the range A + n*(B-A), B + n*(B-A) for appropriate n. Then if NW(I)≠NBC, NXC is stored in NQ(I); otherwise if Y<A, the underflow character NUC is stored in NQ(I), and if Y<B, the overflow character NŌC is stored in NQ(I). If a character is zero, NQ will be left unchanged. If B<A, the computation proceeds as though A and B were interchanged and I was replaced by 101-I.

In addition, if the value of D before rounding differed from an integer by more than 3/8, NGC (or NXC, NUC, or NŌC as appropriate) will be stored in NQ(I+1) or NQ(I-1) so that the two adjacent characters will, when plotted, straddle the proper position. If the value of D before rounding differed from an integer by an amount between 1/8 and 3/8, a decimal point (.) will be stored in NQ(I+1) or NQ(I-1) so that it, along with the graph character NGC stored in NQ(I), will indicate the direction in which the plotted character should have been moved by a slight amount. A decimal point will be stored, however, only if it replaced a blank character NBC.

In use, to plot n functions, GRAPHS would be called n times, with the n function and scale values, then a line would be printed. A suitable print statement might be: WRITE OUTPUT TAPE 61, 99, X, (NQ(I), I=1, 101), 99 FORMAT (1HbE16.9, 2X, 101R1), where X is the abscissa value or perhaps one of the function values.

Restrictions: A must not equal B; zero (0) cannot appear as a plotted character.

USES: No subroutines.

GRAPHS: Plot graphs on the line printer.

FORTRAN FUNCTION
CDC 3800
(Internal to EPAGAS)
IBINSER (X, N, P)

IBINSER
Margot Ackley
July 30, 1973

Given an array X of length N containing monotonically increasing numbers and given a number P, IBINSER performs a binary search in array X to find the integer value I for which $X(I) \leq P > X(I+1)$. IBINSER is then set equal to I, that is, $1 \leq IBINSER \leq N$. If $N=2$, $IBINSER=1$.

USES: No subroutines.

IBINSER: Binary search of a monotonically increasing array.

FORTRAN SUBROUTINE
CDC 3800

(Internal to EPAGAS)
MINMYZD (W, WI, NG, CALC, ERR, N, ZEP, KX, PRNT, RMS, XTM)

MINMYZD
R. Gregg Merrill
(J. R. Winkelman)
March 20, 1970

This subroutine is used to minimize a function of several variables when the true value of the function is known at several points. In other words, it fits an arbitrary function of NG variables to N known points. For example, many positions of a satellite over a few days may be known and the elements (period, major axis, etc.) are desired which will give an orbit for which the rms distance from the given position is a minimum.

Because the convergence procedures and flexibility are much greater than would be required on "nice functions," a suggested set of control factors KX will be given so that the average programmer does not need to become familiar with all the options. (See summary.)

In detail, the subroutine starts with a given set of $NG \leq 10$ parameters W and an increment WI for each W. Within the WI array, a zero implies that the corresponding W is not to be changed. Also given are the $N \leq 160$ errors ERR (an error at each of the given points where these are of no concern to MINMYZD) and an rms error RMS where RMS may be obtained by RMS = RMSERR(ERR,N). NG may be 1, but N must be $\geq NG$.

MINMYZD makes a step in W_i of WI_i and looks at the resultant errors obtained by calling CALC(W,ERR,RMS). If the RMS for this new set is smaller, the slopes of the individual errors with respect to WI_i are calculated and stored. If larger, a step of WI_i is made in the other direction and slopes are calculated. If neither RMS is smaller, WI_i is halved for future use, as the step size is probably too large. If no errors changed at all, WI_i is set to zero (normally because WI was halved out of existence and is no longer important to the minimization process).

After all W's have been varied, the matrix

$$(\partial ERR_i/\partial W_j)*(\Delta W_j) = -(ERR_i) \quad (1)$$

is solved for the changes ΔW_i in W. The new parameters are then $W_i = W_i+ZEP*\Delta W_i$. If the important $\partial ERR_i/\partial W_j$ is nearly constant, ZEP should be 1. As a minimum is approached, ZEP must be made smaller. This is done internally under control of the control values KX.

After the new ERR's and RMS have been calculated, PRNT is checked. If PRNT is zero, no internal printing is done. If PRNT $\neq$ 0, a line is printed specifying the ZEP used and the RMS corresponding to that ZEP. It also prints a ZEP_p and RMS_p which would have theoretically minimized the curve (slopes are known for ZEP = 0, and the theoretical value RMS_p for any ZEP_p can be compared against the actual RMS evaluated at the given ZEP to predict the second derivatives and the corresponding minimum).

If at any time a new RMS is less than the old one, the W and ERR arrays are replaced by the better values and the necessary adjustments made to KX (7 and 8).

The KX array contains six input control parameters and four output parameters. These control the action of MINMYZD. None of the controls KX(1-6) are altered by the program.

KX(1) The total number of cycles to take, even if conditions are still good.

KX(2) The number of cycles to count for a cycle which does not cause enough improvement.

KX(3) The acceptable improvement in 0.1 percent (105 is 10.5 percent). Any cycle resulting in less than this causes KX(2) cycles to be counted rather than one cycle.

KD(4) The limit as to the multiple of a WI increment which may be taken in one full step (ZEP=1). A zero will be used as 100. A small slope might predict a step of 10^6 which is clearly too large. If KX(4) is negative, no step is made in the -WI direction. It is normally faster to have KX(4)<0 for the first few uses of MINMYZD.

KX(5) Minimum ZEP to be used in 0.001 (10 is 0.010). If ZEP is allowed to go too small (alas, the predictions are not always good) the step does not go anywhere. If ZEP minimum is too large, ZEP cannot be made small enough to account for the nonlinearity in ERR for final convergence.

KX(6) If good improvements are made, it is natural for ZEP to become smaller as MINMYZD moves away from the region in which the derivatives were calculated. If the total improvement KX(10) is greater than KX(6) in 0.1 percent, then the ZEP at exit will be the same as the ZEP on entry, that is, things are going well so stick to the same ZEP. Otherwise, exit with the last predicted ZEP_p.

KX(7) through KX(10) are used to convey certain information to the user.

KX(7) If 1, at least one single step (even if it were a setup of WI_i) resulted in an improvement of at least KX(3); otherwise, it is zero.

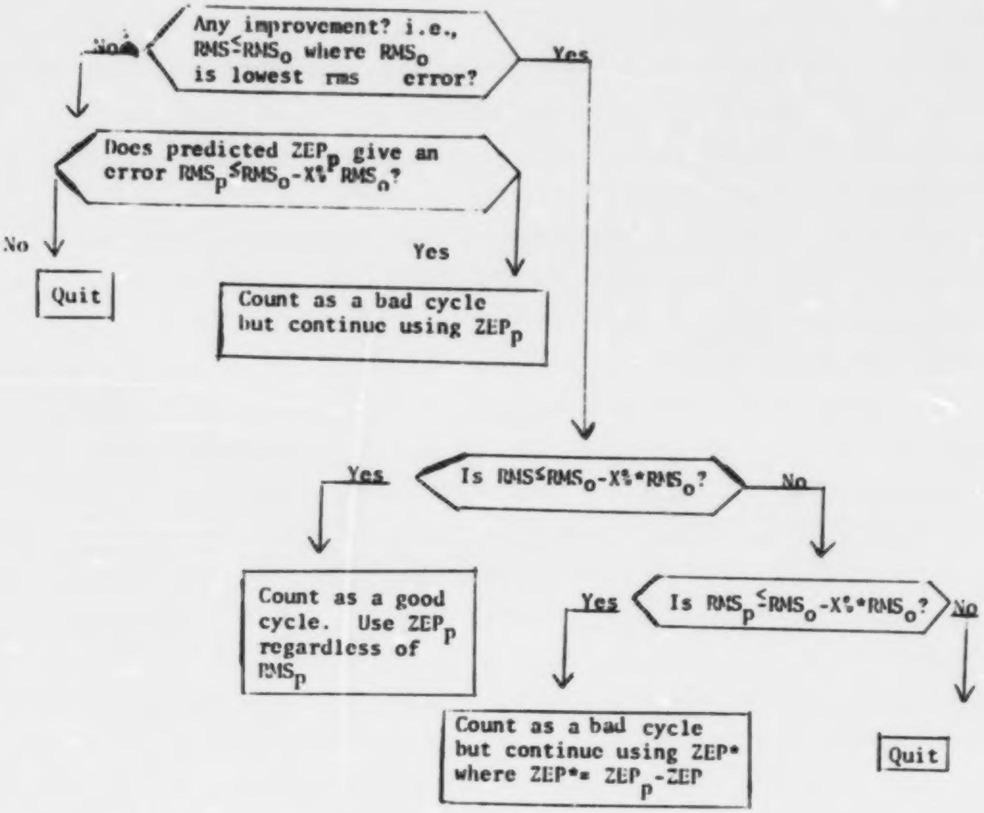
KX(8) If 2, there was some improvement sometime; otherwise, zero.

KX(9) If -1, the WI's are all too small; the system has probably converged. If 0, the matrix was singular; this is not expected to occur. +1 is normal.

KX(10) Total improvement in 0.1 percent relative to the RMS input. 750 would mean the new RMS was ¼ of the old RMS.

After PRNT is checked, the improvement is checked against KX(3). If sufficient, one is subtracted from the cycle count set up from KX(1).

A ZEP gives an error RMS: $(X = 0.1(KX(3)))$


Any improvement? i.e., RMS≤RMS$_o$ where RMS$_o$ is lowest rms error?
No
Yes
Does predicted ZEP$_p$ give an error RMS$_p$≤RMS$_o$-X%*RMS$_o$?
No
Yes
Quit
Count as a bad cycle but continue using ZEP$_p$
Is RMS≤RMS$_o$-X%*RMS$_o$?
Yes
No
Count as a good cycle. Use ZEP$_p$ regardless of RMS$_p$
Is RMS$_p$≤RMS$_o$-X%*RMS$_o$?
Yes
No
Count as a bad cycle but continue using ZEP* where ZEP*= ZEP$_p$-ZEP
Quit

If the count is ≥ 1, the new ERR is put into equation (1) and the system continues from there using the predicted ZEP_p. Any time the cycle count reaches zero, an exit is made.

If the improvement is not good enough, KX(2) is subtracted from the cycle count. If the count ≤ 0 or the predicted RMS_p is also not good enough, an exit is made. Otherwise, a calculation is made using a ZEP based on the old ZEP and the predicted ZEP_p, hoping for an improvement. If no improvement is made but the predicted RMS_p is good enough, then calculations are made using the predicted ZEP_p; otherwise, an exit is made. (See flow chart for termination of MINMYZD.)

Helpful hints: If KX(2) is one less than KX(1), then, after a "good" cycle, one bad cycle will end it. If a "bad" cycle occurs immediately, then one more try with the adjusted ZEP will usually give good results. On complicated functions (almost anything will work on simple ones), it may be necessary to set KX(5) higher from 100 to 300 for a few steps if ZEP is being predicted too small. The XTM matrix, currently (160,11), is the partial derivatives matrix $XTM(I,K) = \partial ERR_i/\partial W_k$, I=1, N, K=1, NG, where K is the Kth non-zero WI_k; that is, there are no zero columns.

If one or more rows of XTM completely dominate, it may result in a singular matrix. This condition is checked, and the matrix is weighted if necessary. The weights are stored in the column (KK) following the normal information. This number XTM(I,KK) will normally be 1. If it is less than 1, then the I^{th} row has been multiplied by XTM(I,KK) and appears smaller than when the row was actually calculated.

SUMMARY:

Given: Arrays W and WI, NG words long where W contains the parameters, WI contains the suggested increments, and a zero WI means that the corresponding W is not to be varied; a subroutine CALC, which given W, calculates an error array ERR (positive and/or negative), N long, and the rms error RMS; a variable ZEP; and values for PRNT, KX(1), KX(2), KX(3), KX(4), KX(5), and KX(6).

MINMYZD exits with improved values of W; WI may be slightly altered; the array ERR of new errors associated with W; RMS correspond to ERR; ZEP may be a better value (must be smaller toward the end); KX(7)=1, at least one single step exceeded 0.1 percent improvement (from KX(3)) or else is zero; KX(8)=1 some improvement occurred sometime or else is zero; KX(9)=1), all is well, =0 unexpected, =-1 all WI's reduced to zero; you are done, congratulations! KX(10) the relative improvement for this one step of MINMYZD is 0.1 percent.

Normally KX(9) should be tested, and then test KX(7,8,or 10) or RMS to determine whether to call MINMYZD again or to quit as good enough. Note that all values needed for input are proper as returned from MINMYZD, so only a call is necessary, no further work.

USES: CALC, CLOSER, internal MINMYZD (FINDIT (PARTIN)), and GAUSSL.

MINMYZD: To minimize an arbitrary function of several variables.

FORTRAN SUBROUTINE
CDC 3800
(Internal to EPAGAS)
PARTIN (K, L, M, N, I, J)

PARTIN
J. R. Winkelman
December 18, 1962

Given the array K, a location in it K(L), and the limits of the area of concern K(M) to K(N), PARTIN shuffles the array until the following inequalities hold:

$$K(IJ) \leq K(L) \qquad \text{for } M \leq IJ \leq J$$
$$K(IJ) \geq K(L) \qquad \text{for } I \leq IJ \leq N$$

where $M \leq J \leq I \leq N$. I and J are output.

PARTIN uses fixed-point arithmetic internally so that it handles fixed and floating arrays without prejudice.

USES: No subroutines.
PARTIN: Partition (partial sort) an array.

FORTRAN SUBROUTINE
CDC 3800
(Internal to EPAGAS)
PLOTD (W, IDGAS)

PLOTD
Margot Ackley
July 30, 1973

Given the above arguments and selected arguments from the two COMMON blocks, COMMON/100/ and COMMON/101/, PLOTD plots an on-line graph of the difference between the given transmissions TG and final computed transmissions TC versus the wavenumber U as a function of the final set of gas concentrations W. Scale for the transmission differences is set from -0.10 to +0.10.

The two COMMON blocks utilized are as follows: COMMON/100/TRAN (5000,5), NU(10), NW(10), UR(10,3), WR(10,5), TG(150), U(150), CF(150), TC(150), WM(10), WHR(10), WW(10), IPW(10); and COMMON/101/NG, NT, RMSTP, RMST, RMSW.

For definitions of arguments and variables used, see below descriptions:

EPAGAS: W
THEOTRN: IDGAS, NG
READ7(8): NT, U, TG
CALC: TC

PLOTD starts a new page.

USES: GRAPHS

PLOTD: Plots difference of given and final computed transmissions versus wavenumber.

FORTRAN SUBROUTINE
CDC 3800
(Internal to EPAGAS)
PLOTT (W, IDGAS, NM)

PLOTT
Margot Ackley
July 30, 1973

Given the above arguments and selected arguments from the two COMMON blocks, COMMON/100/ and COMMON/101/, PLOTT plots an on-line graph of the given transmissions TG and corresponding computed transmissions TC versus the wavenumber U as a function of the current values of the gas concentrations W. Scale for transmissions is set from 0 to 1.

The two COMMON blocks utilized are as follows: COMMON/100/TRAN (5000,5), NU(10), NW(10), UR(10,3), WR(10,5), TG(150), U(150), CF(150), TC(150), WM(10), WHR(10), WW(10), IPW(10); and COMMON/101/NG, NT, RMSTP, RMST, RMSW.

For definitions of arguments and variables used, see below descriptions:

EPAGAS:	W
THEOTRN:	IDGAS, NG
READ7(8):	NT, U, TG
CALC:	TC

If NM=0, a title line stating that the plot is before use of MINMYZD is printed; if NM>0, a title line stating that the plot is after NX uses of MINMYZD is printed; and if NM<0, a title line stating that the plot is after final use of MINMYZD is printed.

PLOTT starts a new page.

USES: GRAPHS

PLOTT: Plot curves of given and currently computed transmission versus wavenumber.

FORTRAN SUBROUTINE
CDC 3800
(Internal to EPAGAS)

PRINT1
Margot Ackley
July 30, 1973

PRINT1 (Q, GLB, CFM, MK, IPRINT, IPLOT, UMIN, UMAX, ZEP, KX, IDGAS, W, WI)

Given all of the above arguments and selected arguments from the two COMMON blocks, COMMON/100/ and COMMON/101/, PRINT1 prints all the input data for program EPAGAS.

The two COMMON blocks are as follows: COMMON/100/TRAN(5000,5), NU(10), NW(10), UR(10,3), WR(10,5), TG(150), U(150), CF(150), TC(150), WM(10), WHR(10), WW(10), IPW(10); and COMMON/101/NG, NT, RMSTP, RMST, RMSW.

For definitions of arguments and variables printed, see below descriptions:

READ4:	Q, GLB, CFM, MK, IPRINT, IPLOT
READ1:	UMIN, UMAX
READ3:	ZEP, KX
THEOTRN:	IDGAS, NG
READALL:	W, WI
READ2:	WM, WHR, WW, IPW
READ7(8):	NT, U, CF, TG

PRINT1 starts a new page.

USES: No subroutines.
PRINT1: Prints all input for EPAGAS.

FORTRAN SUBROUTINE PRINT2
CDC 3800 Margot Ackley
(Internal to EPAGAS) July 30, 1973
PRINT2 (NM, Q, GLB, MK, ZEP, KX, RMS, IDGAS, WO, W, WI, ERR)

Given the above arguments and selected arguments from the two COMMON blocks, COMMON/100/ and COMMON/101/, PRINT2 prints the intermediate data and errors before the first use of MINMYZD and after any successive uses of MINMYZD.

The two COMMON blocks are as follows: COMMON/100/TRAN(5000,5), NU(10), NW(10), UR(10,3), WR(10,5), TG(150), U(150), CF(150), TC(150), WM(10), WHR(10), WW(10), IPW(10); and COMMON/101/NG, NT, RMSTP, RMST, RMSW.

For definitions of arguments and variables printed, see below descriptions:

READ4:	Q, GLB, MK
READ3:	ZEP, KX
CALC:	RMS
THEOTRN:	IDGAS, NG
READ6:	WO
READALL:	W, WI
CALC:	ERR, RMST, RMSW, TC
READ2:	WM, WHR, WW, IPW
READ7(8):	NT, U, CF, TG

If NM=0, a title line stating that the data is before use of MINMYZD is printed. If NM>0, a title line stating that the data is after NM uses of MINMYZD is printed.

PRINT2 starts a new page.

USES: No subroutines.

PRINT2: Prints intermediate data and errors for EPAGAS.

FORTRAN SUBROUTINE
CDC 3800
(Internal to EPAGAS)
PRINT3 (IDGAS, WO, W, ERR, RMS)

PRINT3
Margot Ackley
July 30, 1973

Given the above arguments and selected arguments from the two COMMON blocks, COMMON/100/ and COMMON/101/, PRINT3 prints the final data and errors.

The two COMMON blocks are as follows: COMMON/100/TRAN(5000,5), NU(10), NW(10), UR(10,3), WR(10,5), TG(150), U(150), CF(150), TC(150), WM(10), WHR(10), WW(10), IPW(10); and COMMON/100/NG, NT, RMSTP, RMST, RMSW.

For definitions of the arguments and variables printed, see below descriptions:

THEOTRN:	IDGAS, NG
READ6:	WO
READALL:	W
CALC:	ERR, RMS, RMST, RMSW, TC
READ2:	WM, WHR, WW, IPW
READ7(8):	NT, U, CF, TG

PRINT3 starts a new page.

USES: No subroutines.
PRINT3: Prints final data and errors for EPAGAS.

FORTRAN SUBROUTINE

CDC 3800

(Internal to EPAGAS)

READALL

Margot Ackley

July 30, 1973

READALL (S, UMIN, UMAX, ZEP, KX, Q, GLB, MK, CFM, IPRINT, IPLOT, W, WI, WO, WIO)

Given the array S of length 5 (see CONST description), READALL reads in and checks input data to be used by EPAGAS. Each of the seven different types of data must be preceded by the appropriate header card containing the card type number in column 1.

Card type	Error array element	FORTRAN reading subprogram
1	S(3)	READ1 (UMIN, UMAX)
2	S(4)	READ2 (FAKEARG)
3	none	READ3 (ZEP, KX)
4	none	READ4 (Q, GLB, CFM, MK, IPRINT, IPLOT)
6	S(1)	READ6 (WO, WIO)
7	S(2)	READ7 (UMIN, UMAX, CFM)
8	S(2)	READ8 (UMIN, UMAX, CFM)

S(K) = 0 indicates no error in reading and S(K) ≠ 0 indicates that an error has occurred. If a card type 5 is read in, S(5) = -1 and indicates an error. See above subprogram descriptions for preparation of data. If 1, 2, 3, 4 type data is not read in, then those values given by CONST will be used in computation. Type 6 data must initially be read in, and type 7 or 8 (but not both) must also be initially read in. As many sets of data as desired may be processed. However, a type 9 card must follow every set of data (9 in column 1). This card will cause checking of data and begins computation if no reading errors occur. READALL provides printed error analysis for any errors.

Arrays W and WI are set equal to the arrays WO and WIO. These are the gas concentrations W with corresponding increments WI, and the original concentrations WO with corresponding original increments WIO. If no concentrations are read in after the initial set, then each new set of data will be processed, starting with the original set of concentrations WO and increments of WIO. If one wishes to use the final computed set of concentrations of the last data set as the starting concentrations of the next data set, this may be accomplished by using READ6 with the appropriate header card, followed by a card with a 2 in column 80. This last set of concentrations and increments will then become the original concentrations and increments for all future sets of data.

After the initial set of data, sets of data need only to contain those values which one wants to vary. The unchanged values which will be used in processing the new data set will be those which had last been read in.

READALL utilizes two blocks of COMMON: COMMON/100/TRAN(5000,5), NU(10), NW(10), UR(10,3), WR(10,5), TG(150), U(150), CF(150), TC(150), WM(10), WHR(10), WW(10), IPW(10); and COMMON/101/NG, NT, RMSTP, RMST, RMSW.

USES: READ1, READ2, READ3, READ4, READ6, READ7, and READ8.
READALL: Reads in and checks data for EPAGAS.

FORTRAN FUNCTION
CDC 3800
(Internal to EPAGAS)
READ1 (UMIN, UMAX)

READ1
Margot Ackley
July 30, 1973

READ1 reads in the minimum wavenumber value UMIN and the maximum wavenumber value UMAX for which the given transmissions will range over.

If $UMIN \leq 0$ and/or $UMIN \geq UMAX$, then an error has orrcurred and READ1 returns equal to -1. If no error occurs, then READ1 = 0.

CARD FORMATS:

CONTROL CARD-read by READALL

Col.	Field	Description
1	Integer	Card type - use 1

DATA CARD-read by READ1

Col.	Field	Description
1-10	Floating	UMIN>0
11-20	Floating	UMZX>UMIN

USES: No subroutines.

READ1: Reads in wavenumber range for EPAGAS.

FORTRAN FUNCTION
CDC 3800
(Internal to EPAGAS)
READ2 (FAKEARG)

READ2
Margot Ackley
July 30, 1973

Given NG, the number of gas constituents, READ2 reads in the necessary gas concentration error parameters to be used by EROR1. NG values each of WM, WHR, WW, and IPW are read in.

If WW(K)<0, WM(K)<0, WHR(K)≤0, and/or IPW(K)≤0 for K=1, NG, then a reading error has occurred and READ2 returns equal to -1. If no error occurs, then READ2=0. See EROR1 description for detailed use of WW, WM, WHR, and IPW. FAKEARG is a fake argument and has no real use in the function. It is set to zero.

READ2 utilizes two blocks of COMMON: COMMON/100/TRAN(5000,5), NU(10), NW(10), UR(10,3), WR(10,5), TG(150), U(150), CF(150), TC(150), WM(10), WHR(10), WW(10), IPW(10); and COMMON/101/NG, NT, RMSTP, RMST, RMSW.

CARD FORMATS:

CONTROL CARD-read by READALL

Col.	Field	Description
1	Integer	Card type - use 2

DATA CARD-read by READ 2

1-13	E13.6	WM(K)≥0	
21-33	E13.6	WHR(K)>0	K=1, NG
41-53	E13.6	WW(K)≥0	
69-70	Integer	IPW(K)>0	

USES: No subroutines.

READ2: Reads in gas concentration error parameters.

FORTRAN SUBROUTINE
CDC 3800
(Internal to EPAGAS)
READ3 (ZEP, KX)

READ3
Margot Ackley
June 30, 1973

READ3 reads in the necessary controls needed for subroutine MINMYZD which are ZEP and the first six elements of the array KX.

CARD FORMATS:

CONTROL CARD-read by READALL

Col.	Field	Description
1	Integer	Card type - use 3

DATA CARD-read by READ3

Col.	Field	Description
1-10	Floating	ZEP ($0<ZEP\leq 1$), a number which tells how much of the predicted final step to take.
11-20	Integer	KX(1), the maximum number of new sets of errors (cycles) to use. Must be a whole number. If set to zero, then one cycle is used. If negative, then each increment value used in MINMYZD is decreased by 10 percent after each complete use of MINMYZD, and KX(1) is used for the number of cycles.
21-30	Integer	KX(2), the number of times to count a bad cycle, $KX(2)\geq 1$.

Col.	Field	Description
31-40	Integer	KX(3), the percent improvement in tenths in the rms error to be accepted as a good cycle. Any cycle resulting in less than this causes KX(2) cycles to be counted rather than one cycle, $KX(3) \geq 0$.
41-50	Integer	KX(4), the limit as to the multiple of a WI increment which may be taken in one full step (ZEP=1). If set to zero, then 100 is used. If KX(4) is negative, then no step will be permitted in the -WI direction for the first complete try of MINMYZD. Future tries of MINMYZD, use KX(4).
51-60	Integer	KX(5), minimum ZEP in 0.001's, $KX(5)>0$.
61-70	Integer	KX(6), total percent improvement in tenths, $KX(6)>0$.

USES: No subroutines.

READ3: Reads controls for MINMYZD.

FORTRAN SUBROUTINE — READ4
CDC 3800 — Margot Ackley
(Internal to EPAGAS) — July 30, 1973
READ4 (Q, GLB, CFM, MK, IPRINT, IPLOT)

READ4 reads in the necessary program controls for EPAGAS: Q is the number of percent by which each complete try of MINMYZD must improve the mean square deviation when using CALC, GLB is the minimum rms error permitted before using MINMYZD again, MK is the maximum number of uses of MINMYZD permitted, and CFM is a constant confidence factor to be applied to the given experimental transmission data if individual ones are not provided.

IPRINT=n permits printing of intermediate results using PRINT2 after every nth use of MINMYZD, and IPLOT=m permits a graph of the given and computed transmissions as functions of wavenumber to be plotted after every mth use of MINMYZD. If either IPRINT and/or IPLOT is zero, then no intermediate printing and/or plotting will occur.

CARD FORMATS:

CONTROL CARD-read by READALL

Col.	Field	Description
1	Integer	Card type - use 4

DATA CARD-read by READ4

1-10	Floating	$Q>0$
11-20	Floating	$GLB\geq 0$
21-30	Floating	$CFM>0$
31-40	Integer	$MK>0$
41-50	Integer	$IPRINT\geq 0$
51-60	Integer	$IPLOT\geq 0$

USES: No subroutines.

READ4: Read program controls for EPAGAS.

FORTRAN FUNCTION
CDC 3800
(Internal to EPAGAS)
READ6 (WO, WIO)

READ6
Margot Ackley
July 30, 1973

Given NG, the number of gas constituents, READ6 reads in NG original gas concentrations WO(K) and their corresponding increments WIO(K), K=1, NG. If the first data card read in contains a 2 in column 80 instead of WO(1) and WIO(1), then the original start concentrations and increments will be set to the final concentrations and increments of the previous data run and READ6 is set to -1. If not, then READ6 reads in the rest of WO and WIO values and returns equal to zero.

READ6 utilizes one COMMON block where: COMMON/101/NG, NT, RMSTP, RMST, RMSW.

CARD FORMATS:

CONTROL CARD-read by READALL

Col.	Field	Description
1	Integer	Card type - use 6

DATA CARD-read by READ6

Col.	Field	Description
1-13	E13.6	WO(K)
21-33	E13.6	WIO(K) K=1, NG

USES: No subroutines.

READ6: Reads in starting gas concentrations and corresponding increments.

FORTRAN FUNCTION

CDC 3800

(Internal to EPAGAS)

READ7 (UMIN, UMAX, CFM)

READ7

Margot Ackley

July 30, 1973

Given UMIN, UMAX, and CFM, where UMIN and UMAX are, respectively, the minimum and maximum wavenumbers for the corresponding given transmissions and CFM is a constant confidence factor to be used for each individual transmission value if separate factors are not provided; READ7 reads in individual transmissions TG(K) and corresponding confidence factors CF(K). The transmissions are assumed to be equally spaced in wavenumber starting at UB, stepping in UI increments, and ending at UE. NT number of pairs of TG(K) and CF(K) are read in where NT=1 + (UE-UB)/UT. The first data card (a.) read in gives the values UB, UE, and UI. The following data cards (b.) contain TG and CF.

If any of the following conditions occur: UB<UMIN, UE>UMAX, UI≤0, NT>150, or TG(K)<0, TG(K)>1, CF(K)<0, K=1, NT, then an error has occurred and READ7 returns equal to -1. Otherwise, if no error occurs, then READ7=0 upon return. If CF(K)=0, then CF(K) is set to CFM for all appropriate K=1, NT.

Before return, the array U(K), K=1, NT is filled where U(K)=UB + (K-1)*UI.

READ7 utilizes two blocks of COMMON: COMMON/100/TRAN(5000,5), NU(10), NW(10), UR(10,3), WR(10,5), TG(150), U(150), CF(150), TC(150), WM(10), WHR(10), WW(10), IPW(10); and COMMON/101/NG, NT, RMSTP, RMST, RMSW.

CARD FORMATS:

CONTROL CARD-read by READALL

Col.	Field	Description
1	Integer	Card type - use 7

DATA CARD (a.) - read by READ7

Col.	Field	Description	
1-10	Floating	UB	$UB \geq UMIN$
11-20	Floating	UE	$UE \leq UMAX$
21-30	Floating	UI	$UI>0$

DATA CARD (b.) - read by READ7

Col.	Field	Description	
1-5	F5.4	TG(j)	
6-10	F5.5	CF(j)	
11-15	F5.4	TG(j+1)	
16-20	F5.5	CF(j+1)	j=1, 9, 17, ...<NT
21-25	F5.4	TG(j+2)	
26-30	F5.5	CF(j+2)	
31-35	F5.4	TG(j+3)	
36-40	F5.5	CF(j+3)	
.	.	.	
.	.	.	
.	.	.	
.	.	.	
.	.	.	
71-75	F5.4	TG(j+7)	
76-80	F5.5	CF(j+7)	

USES: No subroutines.

READ7: Reads in given transmissions and corresponding confidence factors for EPAGAS.

FORTRAN FUNCTION
CDC 3800
(Internal to EPAGAS)
READ8 (UMIN, UMAX, CFM)

READ8
Margot Ackley
July 30, 1973

Given UMIN, UMAX, and CFM, where UNIM and UMAX are, respectively, the minimum and maximum wavenumbers for the corresponding given transmissions and CFM is a constant confidence factor to be used for each individual transmission value if separate factors are not provided, READ8 reads in the individual wavenumber values U(K), the corresponding given transmission TG(K), and its confidence factor CF(K), K=1, NT. The first data card (a.) read in gives the value of NT, the number of given transmissions to be read. The following cards (b.) contain TG, U, and CF.

If any of the following conditions U(K)<UMIN, U(K)>UMAX, TG(K)<-0.2, or TG(K)>1.01, or CF(K)<0, K=1, NT, or NT>150 occur, then an error has been encountered and READ8 returns equal to -1. Otherwise, no error occurred and READ returns equal to 0. If CF(K)=0, then CF(K) is set to CFM for all appropriate K=1, NT.

READ8 utilizes two blocks of COMMON: COMMON/101/TRAN(5000,5), NU(10), NW(10), UR(10,3), WR(10,5), TG(150), U(150), CF(150), TC(150), WM(10), WHR(10), WW(10), IPW(10); and COMMON/101/NG, NT, RMSTP, RMST, RMSW.

CARD FORMATS:

CONTROL CARD-read by READALL

Col.	Field	Description
1	Integer	Card type - use 8

DATA CARD (a.) - read by READ8

Col.	Field		Description
1-3	Integer	NT	$0 < NT \leq 150$

DATA CARD (b.) - read by READ8

1-10	Floating	TG(K)	
11-20	Floating	U(K)	K=1, NT
21-30	Floating	CF(K)	

USES: No subroutines.

READ8: Reads in given transmission and corresponding wavenumbers and confidence factors for EPAGAS.

FORTRAN FUNCTION
CDC 3800
(Internal to EPAGAS)
RMSERR (A, N)

RMSERR
J. R. Winkelman
June 15, 1965

Given an array A of length N, RMSERR is the root mean square of the array A where

$$\text{RMSERR} = \left(\frac{1}{N} \sum_{i=1}^{N} A_i^2 \right)^{\frac{1}{2}} .$$

USES: No subroutines.

RMSERR: Root mean square of an array.

FORTRAN FUNCTION — THEOTRN
CDC 3800 — Margot Ackley
(Internal to EPAGAS) — July 30, 1973
THEOTRN (IDGAS)

THEOTRN reads in the gas library of theoretical transmissions for a given number of specified gases as a function of both gas concentration and wavenumber. If any errors occur in reading or certain variables are outside permitted ranges, THEOTRN returns immediately equal to one. If all reading is correct, then THEOTRN returns equal to zero.

Before the actual gas library is read in, THEOTRN first reads in a set of data cards that specify which gases are in the library, which concentrations for each gas appear, and which wavenumber region is used for each gas.

Variables appearing on the data cards with their corresponding dimensions and restrictions, if applicable, are given below:

NG:	Total number of different gases to be found in the library, $0<NG \leq 10$.
IDGAS:	Array of maximum length 10 which contains a numerical (I2) identifier for each of the gases, where IDGAS(I) = identifier for the I^{th} gas, I=1, NG.
NU:	Array of maximum length 10 which contains the number of wavenumber values used for each gas, where NU(I) = number of wavenumber values for the I^{th} gas, I=1, NG.
NW:	Array of maximum length 10 which contains the number of concentrations used for each gas, where NW(I) = number of concentrations for the I^{th} gas, I=1, NG; $0<NW(I) \leq 5$.

UR:	A two-dimensional array UR(I,J) dimensioned (10,3). UR contains the wavenumber range and spacing used for each of the gases. I corresponds to the Ith gas. J=1 corresponds to the beginning wavenumber, J=2 corresponds to the end wavenumber, and J=3 corresponds to the wavenumber increment. Note that 1. + (UR(I,2) - UR(I,1))/UR(I,3) should be equal to NU(I).
WR:	A two-dimensional array WR(I,J) dimensioned (10,5). WR contains the actual concentrations used for each gas. I corresponds to the I^{th} gas. J corresponds to the J^{th} concentration for the particular I^{th} gas and must be given in strictly increasing order where 1.E-20<WR(I,1)<WR(I,2)<...<WR(I,J)<... WR(I,NW(I)).

Note that the number of wavenumber values and the wavenumber range and spacing are the same for all concentrations for a given gas.

DATA CARD FORMATS:

Card 1.

Col.	Field	Description
1-4	Integer	NG

Card 2.

2-4	Integer	IDGAS(1)
6-8	Integer	IDGAS(2)
.	.	.
.	.	.
.	.	.

Col.	Field	Description
.	.	IDGAS(NG)
Card 3.		
1-4	Integer	NU(1)
5-8	Integer	NU(2)
.	.	.
.	.	.
.	.	.
.	.	NU(NG)
Card 4.		
1-4	Integer	NW(1)
5-8	Integer	NW(2)
.	.	.
.	.	.
.	.	.
.	.	NW(NG)
Card 5.		
1-8	Floating	UR(1,1)
9-16	Floating	UR(1,2)
17-24	Floating	UR(1,3)
25-32	Floating	UR(2,1)
33-40	Floating	UR(2,2)
41-48	Floating	UR(2,3)
49-56	Floating	UR(3,1)
57-64	Floating	UR(3,2)
65-72	Floating	UR(3,3)
73-80		Blank

Use successive cards up to a maximum of four.

Card 5. +J, $1 \leq J \leq 13$

Col.	Field	Description
1-10	E10.3	WR(I,1)
11-20	E10.3	WR(I,2) I=1, NG
.	.	.
.	.	.
.	.	.
.	.	WR(I,NW(I))

Use successive NG cards, one for each gas.

THEOTRN next reads in the actual gas library.

LIBRARY CARD FORMAT:

Col.	Field	Description
1-2	Integer	NAME
3-4	Integer	ICON
5-8	Integer	NN
9-12	F4.3	T(J)
13-16	F4.3	T(J+1)
17-20	F4.3	T(J+2)
21-24	F4.3	T(J+3)
.	.	
77-80	F4.3	T(J+17)

The gas with identification IDGAS(1) must come first, then the gas with IDGAS(2) is next, etc. Gas with identification IDGAS(NG) comes last. NAME on each gas library card corresponds to IDGAS of the particular gas being read in. Within a particular gas, the concentrations W must be increasing. The first concentration corresponds to ICON=1, second concentration corresponds to ICON=2, and last concentration corresponds to

ICON=NW(I) for the I^{th} gas. Within a given gas and a given concentration, the cards are serialized by NN where NN=1 for the first card, NN=2 for the second card, etc. The maximum value for NN within a set is NU(I)/18 for the I^{th} gas. T(J) corresponds to the theoretical transmission for the particular Ith gas and its particular concentration. T(1) is the transmission corresponding to first wavenumber UR(I,1), T(2) is the transmission corresponding to the wavenumber UR(K,1) + UR(I,3), and the last transmission, T(NU), corresponds to the last wavenumber UR(I,2) for the I^{th} gas. Within a given gas and particular concentration, all cards have 18 values of transmission except for the last card which may have less.

THEOTRN does extensive checking of the card order when reading in the library. If the cards are out of order, an error has occurred and THEOTRN terminates equal to one.

As the individual transmission T is read in, it is stored in the large transmission array TRAN (I,J) which has been dimensioned (5000,5). J corresponds to the concentration number ICON, where $1 \leq J \leq 5$ for all gases. I=1 corresponds to the transmission of the first gas, first wavenumber. I=NU(1) corresponds to the transmission of the first gas, last wavenumber. I=NU(1) + 1 corresponds to the transmission of the second gas, first wavenumber. I = NU(1) + NU(2) corresponds to the transmission of the second gas, last wavenumber. $I = \sum_{i=1}^{NG} NU(i) \leq 5000$ corresponds to the transmission of the last gas, last wavenumber. Unless all gases have the same number of concentrations, the actual amount of storage utilized will not be rectangular.

THEOTRN utilizes two COMMON blocks COMMON/100/ and COMMON/101/ where COMMON/100/TRAN(5000,5), NU(10), NW(10), UR(10,3), WR(10,5), TG(150), U(150), CF(150), TC(150), WM(10), WHR(10), WW(10), IPW(10); and COMMON/101/NG, NT, RMSTP, RMST, RMSW.

USES: No subroutines.

THEOTRN: Reads in gas library of theoretical transmissions for EPAGAS.

A2.2 SIMULATION OF GAUSSIAN NOISE

To study the effects of noise in the transmission data, a method of obtaining random Gaussian noise was developed. Many large-scale computers (CDC 3800 in particular) provide a library routine which furnishes the user with a sequence of uniformly distributed random numbers ranging from 0 to 1. Given two such numbers A_i and B_i, a Gaussian random variable X_i can be obtained from a transformation provided by Box and Muller (1958), where

$$X_i = (-2 \log_e A_i)^{\frac{1}{2}} \cos (2\Pi B_i).$$

Three different signal-to-noise ratios 10, 30, and 100 were incorporated in this report. Letting SNR be the signal-to-noise ratio desired, then SNR is defined as

$$\mathrm{SNR} = \frac{\sigma_d}{\sigma_n} \quad ,$$

where σ_d is the standard deviation of the data and σ_n is the standard deviation of the Gaussian noise for the set of data to be analyzed. The noise N_i used to modify a particular transmission value T_i is defined as

$$N_i = \sigma_n X_i \ ,$$

where the noise over the set has mean zero and standard deviation σ_n.

The new transmission values TN_i incorporating the noise are then defined as

$$TN_i = T_i + N_i \quad .$$

A2.3 CONSTITUENT AND TRANSMISSION ERROR FORMULATION

To direct its progress, the minimization process uses a total rms error E as a measure of the "goodness of fit." By varying the amounts of each of the gas constituents, the process tries to minimize the total rms error E which is both a function of the rms error E_T from the individual transmission errors and the rms error E_W from the errors of the individual constituent amounts. (See EROR2 and EROR1 descriptions, respectively.)

For the benefit of the user, all three rms errors E, E_T, and E_W are standard printed variables. In this program, the total rms error E and the constituent rms error E_W are calculated. From these two, the transmission rms error E_T can then be calculated as follows.

Let NG = the number of gas constituents,
NT = the number of transmission values,
A_i = the individual errors from the gas constituents, i=1, NG,
A_i = the individual errors from the transmissions, i=NG+1, NG+NT,
N = NG + NT.

Then,

$$E = \left[\frac{1}{N} * \sum_{i=1}^{N} A_i^2 \right]^{1/2}$$

$$E_W = \left[\frac{1}{NG} * \sum_{i=1}^{NG} A_i^2 \right]^{1/2}$$

But

$$E^2 = \frac{1}{N} * \sum_{i=1}^{NG} A_i^2 + \frac{1}{N} * \sum_{i=NG+1}^{NG+NT=N} A_i^2$$

$$E^2 = \frac{NG}{N} * \frac{1}{NG} * \sum_{i=1}^{NG} A_i^2 + \frac{NT}{N} * \frac{1}{NT} * \sum_{i=NG+1}^{N} A_i^2$$

and

$$E_T^2 = \frac{1}{NT} * \sum_{i=NG+1}^{N} A_i^2 \quad .$$

Therefore,

$$E^2 = \frac{NG}{N} * E_W^2 + \frac{NT}{N} * E_T^2$$

$$E_T = \left[\frac{N * E^2 - NG * E_W^2}{NT} \right]^{\frac{1}{2}} \quad .$$

A2.4 INTERPOLATION FORMULAS USED IN TRANSMISSION LIBRARY

Given a wavenumber ν and a gas concentration W for a specified constituent, a corresponding transmission $T(\nu,W)$ can be found from the transmission library. Because the library contains transmissions at discrete values of wavenumber and concentration, a form of double linear interpolation must be utilized.

Assuming the ν and W lie within the region of the transmission table, that is,

$$W_{min} \le W_1 \le W \le W_2 \le W_{max}$$

$$\nu_{min} \le \nu_1 \le \nu \le \nu_2 \le \nu_{max} ,$$

then by interpolating first in the W direction, we have,

$$\frac{T(\nu_1,W) - T(\nu_1,W_1)}{W - W_1} = \frac{T(\nu_1,W_2) - T(\nu_1,W_1)}{W_2 - W_1} \qquad \text{(A2:1)}$$

and

$$\frac{T(\nu_2,W) - T(\nu_2,W_1)}{W - W_1} = \frac{T(\nu_2,W_2) - T(\nu_2,W_1)}{W_2-W_1} .$$

Therefore,

$$T(\nu_1,W) = \frac{W-W_1}{W_2-W_1} * \left[T(\nu_1,W_2) - T(\nu_1,W_1) \right] + T(\nu_1,W_1),$$

and

$$T(\nu_2,W) = \frac{W-W_1}{W_2-W_2} * \left[T(\nu_2,W_2) - T(\nu_2,W_1) \right] + T(\nu_2,W_1). \qquad \text{(A2:2)}$$

Interpolating next in the ν direction, we have

$$\frac{T(\nu,W) - T(\nu_1,W)}{\nu - \nu_1} = \frac{T(\nu_2,W) - T(\nu_1,W)}{\nu_2 - \nu_1} . \qquad \text{(A2:3)}$$

Therefore,

$$T(\nu,W) = \frac{\nu-\nu_1}{\nu_2-\nu_1} * \Big[T(\nu_2,W) - T(\nu_1,W) \Big] + T(\nu_1,W). \qquad \text{(A2:4)}$$

By substituting the expressions for $T(\nu_1,W)$ and $T(\nu_2,W)$ from (A2:2) into equation (A2:4), we have the following interpolation formula

$$T(\nu,W) = \frac{\nu-\nu_1}{\nu_2-\nu_1} * \left[\frac{W-W_1}{W_2-W_1} * \Big(T(\nu_2,W_2) - T(\nu_2,W_1) + T(\nu_1,W_1) - T(\nu_1,W_2) \Big) + T(\nu_2,W_1) - T(\nu_1,W_1) \right] + \frac{W-W_1}{W_2-W_1} * \Big[T(\nu_1,W_2) - T(\nu_1,W_1) \Big] + T(\nu_1,W_1). \qquad \text{(A2:5)}$$

Because of the structure of the program, the wavenumber ν will always lie within the region of the transmission table. However, the concentration W for a particular gas constituent could possibly lie outside the region due to the minimization process which varies only the W's. In the event W is outside the region, we have two possible cases:

(1) $W_{max} < W$

(2) $W < W_{min}$.

For the first case, let

$$W_{max-1} = W_1 < W_{max} = W_2 < W.$$

Then,

$$\frac{T(\nu_1,W) - T(\nu_1,W_1)}{W-W_1} = \frac{T(\nu_1,W_2) - T(\nu_1,W_1)}{W_2-W_1}$$

$$\frac{T(\nu_2,W) - T(\nu_2,W_1)}{W-W_1} = \frac{T(\nu_2,W_2) - T(\nu_2,W_1)}{W_2-W_1} , \qquad (A2:6)$$

and

$$\frac{T(\nu,W) - T(\nu_1,W)}{\nu- \nu_1} = \frac{T(\nu_2,W) - T(\nu_1,W)}{\nu_2 - \nu_1} . \qquad (A2:7)$$

By referring to (A2:1) and (A2:3), we can then use (A2:5) for the final formula with $W_1 = W_{max-1}$ and $W_2 = W_{max}$.

For the second case, where

$$W_1 = 0 \leq W < W_{min} = W_2$$

or

$$W_1 = 0 < W_{min} = W_2 ,$$

we may utilize the fact that

$$T(\nu,0) = T(\nu,W_1) \equiv 1 \text{ for all } \nu. \qquad (A2:8)$$

Applying (A2:8) to (A2:5), we have

$$T(\nu,W) = \frac{\nu-\nu_1}{\nu_2-\nu_1} * \left[\frac{W-0}{W_2-0} * \left(T(\nu_2,W_2)-1+1-T(\nu_1,W_2) \right) + 1-1 \right] +$$

$$\frac{W-0}{W_2-0} * \left[T(\nu_1,W_2) - 1 \right] + 1. \qquad (A2:9)$$

Therefore,

$$T(\nu,W) = 1 + \frac{W}{W_2} * \left[\frac{\nu-\nu_1}{\nu_2-\nu_1} * \left(T(\nu_2,W_2) - T(\nu_1,W_2)\right) + T(\nu_1,W_2) - 1\right], \qquad \text{(A2:10)}$$

where

$$W_2 = W_{min}.$$

In summary, given ν and W to find $T(\nu,W)$, where $\nu_{min} \le \nu_1 \le \nu \le \nu_2 \le \nu_{max}$:

For $W_{min} \le W_1 \le W \le W_2 \le W_{max}$,

$$T(\nu,W) = \frac{\nu-\nu_1}{\nu_2-\nu_1} * \left[\frac{W-W_1}{W_2-W_1} * \left(T(\nu_2,W_2) - T(\nu_2,W_1) + T(\nu_1,W_1) - T(\nu_1,W_2)\right) + T(\nu_2,W_1) - T(\nu_1,W_1)\right] + \frac{W-W_1}{W_2-W_1} * \left[T(\nu_1,W_2) - T(\nu_1,W_1)\right] + T(\nu_1,W_1).$$

For $W_{max} < W$,

$$T(\nu,W) = \frac{\nu-\nu_1}{\nu_2-\nu_1} * \left[\frac{W-W_{max-1}}{W_{max}-W_{max-1}} * \left(T(\nu_2,W_{max}) - T(\nu_2-W_{max-1}) + T(\nu_1,W_{max-1}) - T(\nu_1,W_{max})\right) + T(\nu_2,W_{max-1}) - T(\nu_1,W_{max-1})\right] + \frac{W-W_{max-1}}{W_{max}-W_{max-1}} * \left[T(\nu_1,W_{max}) - T(\nu_1,W_{max-1})\right] + T(\nu_1,W_{max-1}).$$

For $W < W_{min}$,

$$T(\nu,W) = 1 + \frac{W}{W_{min}} * \left[\frac{\nu-\nu_1}{\nu_2-\nu_1} * \left(T(\nu_2,W_{min}) - T(\nu_1,W_{min})\right) + T(\nu_1,W_{min}) - 1.\right]$$

A2.5 PROGRAM EXECUTION TECHNIQUES

Although program EPAGAS is a totally self-contained program, its proper utilization requires art as well as science if successful results are desired. Before attempting a run, the user should thoroughly acquaint himself with all the program and subprogram descriptions. The following categories require careful checking before the program is run and are a trouble-shooting checklist in case of improper results.

A2.5.1 Theoretical Gas Library Spectra

The user cannot hope to have good results if the library of spectra does not contain all of the gas constituents which are (or are suspected) to occur in the experimental data. Also the wavenumber range for the gas constituents may not be sufficient for the range of the experimental data. If this occurs, an error message will be printed. The immediate solution to this problem is either to increase the wavenumber range of the library spectra or to analyze a smaller region of the experimental data.

The range of the concentrations of a particular gas constituent is another item that should be checked. The library should have a range of concentrations that covers the expected concentration of a particular gas constituent. If, during the process of trying to find a solution, EPAGAS tries to use a concentration outside the region of given concentrations, a method of linear extrapolation is used to derive a particular gas concentration. For concentrations less than 10% of the total range of the region lying beyond the outer limits of the library, there generally will not be problems. However, concentrations greater than 10% could cause the program to give erroneous results.

The spacing of the library spectra in both wavenumber and concentration is also very important. The program uses linear interpolation in both wavenumber and concentration to determine transmissions for a given wavenumber and a trial concentration. (See FINDTRN decription.) If the grid spacing is not fine enough to give reasonably correct results with linear interpolation, the program may not be able to give satisfactory results. The fineness of the library is often limited by available computer memory and by the total number of gas constituents present. In the case where the user has experimental data containing many gas constituents, it might be necessary to limit the wavenumber region to a fairly narrow band or to pick regions where only a few gas constituents are prominent.

The program has extensive built-in checks to make sure that the library spectra and associated wavenumber and concentration information are properly read in. If any of the library cards are out of order or if inconsistencies occur, an error message is printed and further execution is terminated. In the event this happens, carefully check the THEOTRN description for proper input of the library spectra.

A2.5.2 Experimental Transmission Data

Unsuccessful results can often be caused by experimental data which have an intolerable noise level. Signal-to-noise ratio, S/N, should be on the order of 100 or better. For S/N<100, the user should select wavenumber regions that have only a single gas constituent or regions which have little interference from other gases.

Before utilizing EPAGAS, the user may find that it is necessary to do some "preprocessing" of the data. Due to a variety of reasons, instrumental drift in particular, the data may have to be corrected either linearly or nonlinearly for both wavenumber and/or transmission. Absolute transmission values of I/I_0 without noise should range from 0 to 1. Past experience in the use of program EPAGAS has shown that wavenumber drift is particularly detrimental to obtaining successful results. Wavenumber and intensity calibrations can be obtained from careful comparison with known spectra.

Another very important item to check is the instrumental resolution. Resolution should be the same or very nearly the same as the resolution used in producing the library spectra. It is recommended that instrumental resolution should deviate from that of the library by no more than 10 percent. Also the resolution should not fluctuate as a function of wavenumber which can be a major problem for some types of spectrometers unless it is corrected.

The temperature and pressure of the experimental data should also correspond as close as possible to those used in producing the library spectra.

In order that the experimental data is well defined, spacing in wavenumber should be sufficiently fine. If the spacing is too coarse, important spectral line features may be lost. If this is the case, a suitable solution would almost be impossible to find. Wavenumber spacings of at least 1/5 of the resolution are strongly recommended.

Two other important aspects of the experimental data should also be taken into consideration. The slit function of the instrument should closely approximate that of the library. The digital resolution in the conversion of the analog I/I_o information to digital form has a definite effect on the final accuracy of determinations of gas concentrations. The higher the digital resolution, the more accurately EPAGAS can determine the concentrations.

A2.5.3 Minimization or "Goodness of Fit" Controls

Often less than desirable results are obtained because EPAGAS has not been permitted to "work hard enough" at finding a satisfactory solution. A thorough understanding of the descriptions of MINMYZD, READ3, READ4, and CXMIZ would be extremely helpful (along with some experience in using EPAGAS) in remedying this problem. MINMYZD description gives the main minimization procedures and presents a discussion of the convergence controls that direct its progress toward a solution. READ3 description provides information on the controls, used by MINMYZD, which are labelled ZEP and array KX. Particular attention should be paid to the first six elements of the array KX. The description of READ4 gives a discussion of the variables Q, GLB, and MK which are used in conjunction with the 9th and 10th elements of the array KX by CXMIZ to determine if another try by MINMYZD to find a solution should be made. It is recommended that when first using the program on a new set of experimental data, the controls should be set as loosely as computer costs will permit.

A2.5.4 Error Parameters

Proper formulation of the errors defined on the individual gas concentrations is most helpful for successful utilization of EPAGAS. The user though must keep in mind that the term "error" is used rather loosely in connection with the gas concentrations. These errors are mainly a way in which to guarantee that the minimization process does not try to select individual gas concentrations which are known to be physically unsuitable. The errors (or better, restraints) on the individual gas concentrations are of the general form,

$$ERR(I) = WW(I)*\left(\frac{W(I)-WM(I)}{WHR(I)}\right)^{IPW(I)} ,$$

where I = i^{th} gas constituent. W is the current trial concentration being used by the minimization process, and all the other variables are input error parameters. (See EROR1 description.) WM is considered to be the mean acceptable concentration. For most cases (unless additional knowledge exists of the amount of the gas constituent present), WM is simply the mean concentration value based on the highest and lowest library concentrations of the particular gas constituent. WHR is the half range of permitted variation of the particular concentration and is simply the difference between the highest library concentration and the mean concentration. As long as W remains within the concentration range of the library, the absolute value of $\left|\frac{W-WM}{WHR}\right|$ will be between 0 and 1, resulting in a fairly small error. If the minimization process tries to select W's outside of the library region, the error will become rapidly larger as W gets farther and farther away from the acceptable region. The variable IPW serves the purpose of determining just how fast the errors will grow as W moves farther and farther away. IPW should always be an odd, positive number to preserve the "direction" of the error. The larger IPW is, the faster the error will increase once W has entered the unacceptable regions. From past experience, IPW=3 usually is a good starting value. WW is a weighting or normalization factor and is used to keep individual gas constituent errors in line with each other and also in proportion with the individual transmission errors. Sometimes several runs have to be made before a proper determination of the WW's can be made; however, unity is always good for an initial try.

It is generally best to keep the gas constituent error parameters fairly loose; however, if prior knowledge of the concentration of a particular gas constituent in the experimental data exists, then one may wish to narrow the acceptable region around the suspected concentration. This would be done by setting WM equal to the suspected concentration, by considerably decreasing WHR, by increasing IPW up to 9 or 11, and by setting WW to 10 or 20 times the values of WW used for the other gas constituents.

The errors, depending on the computed and given transmission values, are of the general form,

$$\mathrm{ERR(I) = CF(I)*(TG(I)-TC(I))},$$

where I = i^{th} value corresponding to the i^{th} wavenumber. TG is the given experimental transmission, TC is the currently computed transmission based on the current gas concentrations as determined from the minimization process, and CF is a confidence or weighting factor. The role of CF is twofold. If some regions of the experimental data are more reliable than others, then the corresponding CF's should be larger than those applied to transmissions in the more uncertain regions. This has an effect on the minimization process so as to let it somewhat "ignore" the less reliable data and not to spend its efforts seeking some erroneous solution to very noisy or bad data. Also, the CF's play a similar role for the transmission errors that the WW's do for the gas concentration errors. Care should be taken to select the CF's so that the transmission errors are proportional to the concentration error. The user must always keep in mind that the minimization process does not involve the individual gas concentration or transmission errors, but rather tries to minimize the composite rms error of all the errors. Hence, it is vitally important that the effects of the gas concentration errors and the transmission errors on the rms error are properly balanced.

A2.5.5 Miscellaneous Execution Techniques

To commence its minimization process, program EPAGAS needs to be furnished with a set of beginning gas constituent concentrations. If no prior knowledge of the concentrations exists, either the mean or the center concentration is always a good choice. As mentioned in the previous section, if the concentration of a particular gas constituent is approximately known, then this value should be the one selected for the beginning gas concentration. Also, in this case, the gas error parameters should be tightened for this particular gas constituent. From prior experience, improper selection of beginning trial gas concentrations has generally had no disastrous effects on the final solution except to increase computation time. In most cases, a unique solution is found regardless of the starting point. However, in some rare cases, a change in the starting positions or a change in the order of the gas constituents can sometimes be helpful in seeking a suitable solution. The size of the initial increments by which the individual concentrations are varied by the minimization process must also be carefully selected. If the increments are too large, the first change in a particular gas concentration may completely "jump" over the actual true concentration. On the other hand, if the increments are too small, a proper solution may not be reached because applying the increment to the original gas concentration results in an insignificant change. When in doubt, use an increment on the larger size because the minimization process will decrease the size of the increment, but will never increase it.

Program EPAGAS is equipped to furnish intermediate printing of current gas concentrations, of currently computed transmissions, and of corresponding errors, along with a graph showing the given and currently computed transmissions as a function of wavenumber. Careful studying of these intermediate results can often furnish the user with a better feeling of how certain changes in the individual gas concentrations affect the overall rms error.

The intermediate printing also provides the user with the current values of selected elements of the KX array, KX(7) through KX(10). These values also can give the user an insight into the success of each cycle or try of the minimization process. (See MINMYZD description.)

Last, but not least, there are some sets of experimental data which do not lend themselves properly to the minimization technique of determining atmospheric gas constituents and their corresponding concentrations from their superimposed infrared spectra. This may occur whenever the data is too noisy or too imprecise, or whenever the scales (I/I_0 or ν) must be corrected nonlinearly. Without good data, good results cannot be achieved, but if the nonlinear corrections are known, they may be applied in the computer program.

APPENDIX 3

TABLE OF CONTENTS

Page

A3. Characteristics of the ROSE Long-Path Spectrophotometer System 185

APPENDIX 3

A3. CHARACTERISTICS OF THE ROSE LONG-PATH SPECTROPHOTOMETER SYSTEM

Design range	(transmission) 0.4 to 4.0 km (0.25 to 2.5 mi.)
Spectral regions	3 to 5.5 μ (1820 to 3330 cm^{-1} and 7 to 13.5 μ (740 to 1430 cm^{-1})
Resolution	to 0.01 μ
Scan time	2-min. minimum (either spectral region), 130-min. maximum (either spectral region)
Display	wavenumber, I, I/I_0 (see section 1.1)
Source and Reference Blackbodies	
Operating temperature	1800°K
Temperature stability	±4°C
Emissivity	0.99 ± 0.01
Aperture	12.7 x 6.3 mm (0.50 x 0.25 in.)
Telescopes	
Focal length	304.8 cm (120 in.)
Clear aperture	61.0 cm (24 in.)
Field of view	0.25 deg. (±0.125 deg.)
Blur circle diameter	150 μ
Primary f/no	2.0
Secondary magnification	2.5
Source Modulator	
Operating frequency	570 Hz
Reference Chopper	
Frequency	330 Hz
Stability	±0.3 Hz

Monochromator

Mfgr. and model	Perkin Elmer (Norwalk, CT) Model 210B
Type	Littrow grating with linear wavenumber drive
Gratings	3 to 5.5 μ; 240 l/mm 20° 3400 cm^{-1} (2.94 μ) 7 to 13.5 μ; 101 l/mm 22°2' 1333 cm^{-1} (7.5 μ) masked to 15.0 cm^{2} net area (nominal)
Slits	width, 0 to 2.0 mm micrometer-controlled; height, 0 to 12.0 mm micrometer-controlled
Wavenumber drive	3/32 to 6 rpm in steps of 2x
Min. scan time	approximately 2 minutes

Detector

Mfgr. and model	Santa Barbara Research Center (Goleta, CA) 8679^{-1}
Type	Mercury-doped Germanium
Operating temp.	26°K (58 psia H_2)
Size	0.2 x 2.0 mm (0.008 x 0.08 in.)
Field of view	90°
D* (λm, 630)	0.84 x 10^{10}m Hz$^{1/2}$/W

Detector Cooler

Mfgr. and model	Cryogenic Technology Inc. (Waltham, MA) Model 20 with 3833-005 temperature controller
Type	closed-cycle He refrigerator, consisting of compressor and cold head connected by flexible hoses
Operating temp. range	approximately 10° to 28°K
Operating temperature	26°K (58 psia H_2)
Cooldown time	approximately 15 min.

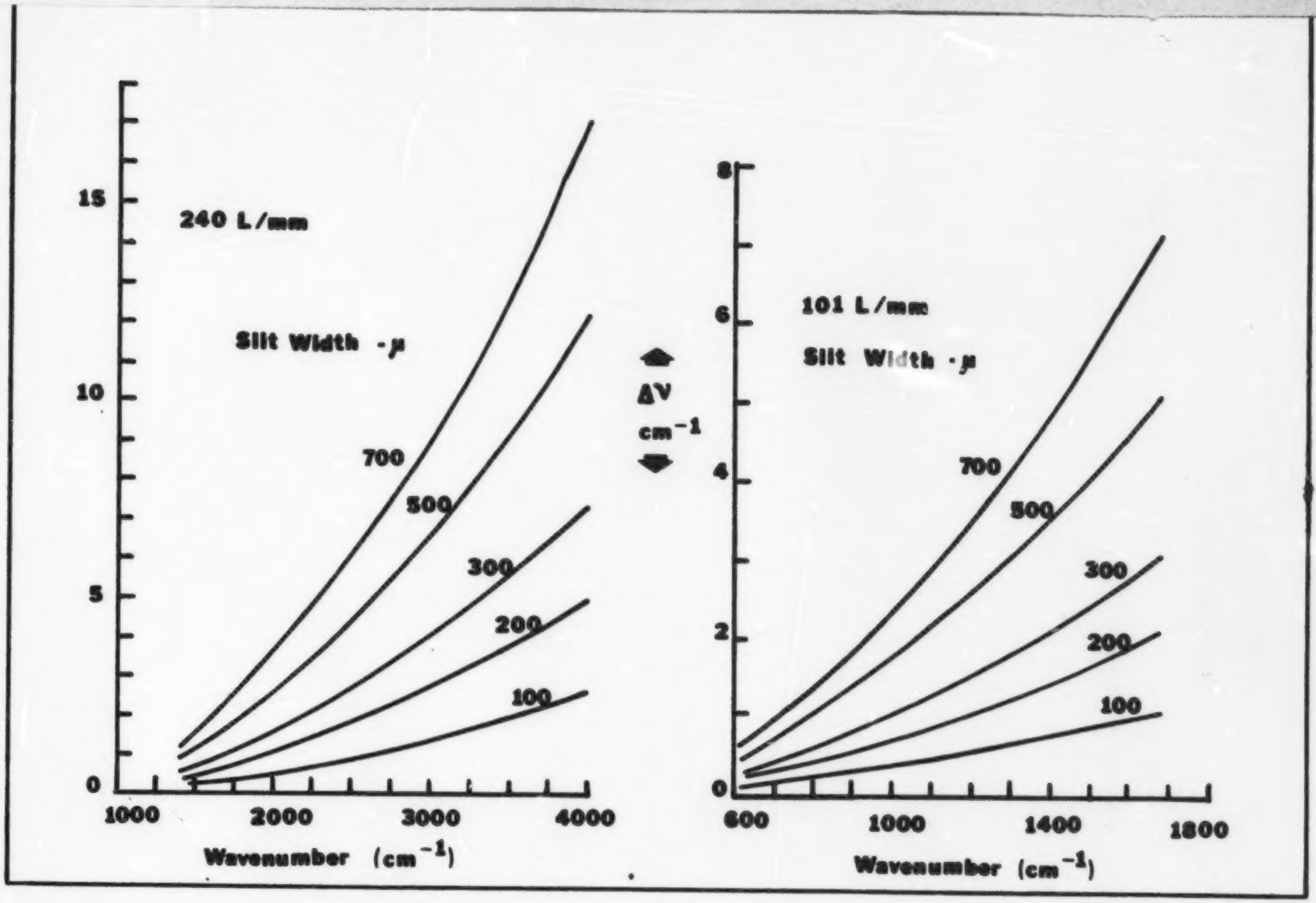


Figure A3.1

RESOLUTION VS. WAVENUMBER FOR VARIOUS SLIT WIDTH SETTINGS
FOR ROSE SPECTROPHOTOMETER

APPENDIX 4

TABLE OF CONTENTS

Page

A4. Convenient Conversion Factors 189

APPENDIX 4

A4. CONVENIENT CONVERSION FACTORS

U (atm cm) STP	X	$1.219 \times 10^{-2} \times \frac{M}{\Theta}$	=	gm cm^{-2}
U (gm cm^{-2})	X	$82.06 \times \frac{273}{M}$	=	(atm cm) STP
U (atm cm) STP	X	2.689×10^{19}	=	molecules cm^{-2}
U (gm cm^{-2})H_2O	X	3.34×10^{22}	=	molecules cm^{-2}
K(ν)(atm cm)$^{-1}$ STP	X	$82.06 \times \frac{\Theta}{M}$	=	(gm cm^{-2})$^{-1}$
K(ν)(gm cm^{-2})$^{-1}$	X	$1.219 \times 10^{-2} \times \frac{M}{273}$	=	(atm cm)$^{-1}$ STP
K(ν)(atm cm)$^{-1}$ STP	X	3.72×10^{-20}	=	(molecules cm^{-2})$^{-1}$
K(ν)(gm cm^{-2})$^{-1}$	X	$\frac{A}{M}$	=	(molecules cm^{-2})$^{-1}$
K(ν)(atm cm)$^{-1}$ STP	X	$356.3 \times \frac{\Theta}{M}$	=	db/(gm cm^{-2})
K(ν)(atm cm)$^{-1}$ STP	X	4.343	=	db/(atm cm) STP
K(ν)(gm cm^{-2})$^{-1}$	X	4.343	=	db/(gm cm^{-2})
$S\left(\frac{cm^{-1}}{\text{atm cm}}\right)$ STP	X	$82.06 \times \frac{\Theta}{M}$	=	$\left(\frac{cm^{-1}}{\text{gm } cm^{-2}}\right)$
$S\left(\frac{cm^{-1}}{\text{gm } cm^{-2}}\right)$	X	$1.219 \times 10^{-2} \times \frac{M}{273}$	=	$\left(\frac{cm^{-1}}{\text{atm cm}}\right)$ STP
$S\left(\frac{cm^{-1}}{\text{atm cm}}\right)$ STP	X	3.72×10^{-20}	=	$\left(\frac{cm^{-1}}{(\text{molecules } cm^{-2})}\right)$
$S\left(\frac{cm^{-1}}{\text{gm } cm^{-2}}\right)$	X	$\frac{M}{A}$	=	$\left(\frac{cm^{-1}}{(\text{molecules } cm^{-2})}\right)$
ppm	X	44.96 M	=	μg m^{-3}

Θ = temperature deg. K,

M = molecular weight

A = Avogadro's number = 6.0225×10^{23}

S = Line intensity

K = Absorption coefficient

U = Optical path